V. S. Varadarajan

Geometry of Quantum Theory

Second Edition

Springer-Verlag
New York Berlin Heidelberg Tokyo

C04234533

QC
174.1
.V35
1985

P756125

LIBRARY
The University of Texas
at San Antonio

WITHDRAWN
UTSA LIBRARIES

Geometry of Quantum Theory

V. S. Varadarajan

Geometry of Quantum Theory

Second Edition

Springer-Verlag
New York Berlin Heidelberg Tokyo

V. S. Varadarajan
Department of Mathematics
University of California
Los Angeles, CA 90024
U.S.A.

AMS Classification: 81-01

Library of Congress Cataloging in Publication Data
Varadarajan, V. S.
Geometry of quantum theory.
Bibliography: p.
Includes index.
1. Quantum theory. I. Title.
QC174.12.V35 1985 530.1′2′01516 85-2738

With 5 Illustrations.

The first edition of this book was published in two volumes: Volume I in 1968 by D. Van Nostrand Company, Inc., New York; and Volume II in 1970 by Van Nostrand Reinhold Company, New York.

© 1968, 1970, 1985 by Springer-Verlag New York Inc.
All rights reserved. No part of this book may be translated or reproduced in any form without written permission from Springer-Verlag, 175 Fifth Avenue, New York, New York 10010, U.S.A.

Printed and bound by R. R. Donnelly & Sons, Harrisonburg, Virginia.
Printed in the United States of America.

9 8 7 6 5 4 3 2 1

ISBN 0-387-96124-0 Springer-Verlag New York Berlin Heidelberg Tokyo
ISBN 3-540-96124-0 Springer-Verlag Berlin Heidelberg New York Tokyo

LIBRARY
The University of Texas

TO MY PARENTS

Digitized by the Internet Archive
in 2023 with funding from
No Sponsor

https://archive.org/details/geometryofquantu0002vara

PREFACE TO VOLUME I OF THE FIRST EDITION

The present work is the first volume of a substantially enlarged version of the mimeographed notes of a course of lectures first given by me in the Indian Statistical Institute, Calcutta, India, during 1964-65. When it was suggested that these lectures be developed into a book, I readily agreed and took the opportunity to extend the scope of the material covered.

No background in physics is in principle necessary for understanding the essential ideas in this work. However, a high degree of mathematical maturity is certainly indispensable. It is safe to say that I aim at an audience composed of professional mathematicians, advanced graduate students, and, hopefully, the rapidly increasing group of mathematical physicists who are attracted to fundamental mathematical questions.

Over the years, the mathematics of quantum theory has become more abstract and, consequently, simpler. Hilbert spaces have been used from the very beginning and, after Weyl and Wigner, group representations have come in conclusively. Recent discoveries seem to indicate that the role of group representations is destined for further expansion, not to speak of the impact of the theory of several complex variables and function-space analysis. But all of this pertains to the world of interacting subatomic particles; the more modest view of the microscopic world presented in this book requires somewhat less. The reader with a knowledge of abstract integration, Hilbert space theory, and topological groups will find the going easy.

Part of the work which went into the writing of this book was supported by the National Science Foundation Grant No. GP-5224. I have profited greatly from conversations with many friends and colleagues at various institutions. To all of them, especially to R. Arens, R. J. Blattner, R. Ranga Rao, K. R. Parthasarathy, and S. R. S. Varadhan, my sincere thanks. I want to record my deep thanks to my colleague Don Babbitt who read through the manuscript carefully, discovered many mistakes, and was responsible for significant improvement of the manuscript. My apologies are due to all those whose work has been ignored or, possibly, incorrectly (and/or insufficiently) discussed. Finally, I want to acknowledge that this book might never have made its way into print but for my wife. She typed the entire manuscript, encouraged me when my enthusiasm went down, and made me understand some of the meaning of our ancient words,

कर्मण्येवाधिकारस्ते मा फलेषु कदाचन। .*

To her my deep gratitude.

Spring, 1968 V. S. VARADARAJAN

---

* *Bhagavadgita, 2:47a.*

PREFACE TO THE SECOND EDITION

कर्मण्येवाधिकारस्ते मा फलेषु कदाचन ।*

It was about four years ago that Springer-Verlag suggested that a revised edition in a single volume of my two-volume work may be worthwhile. I agreed enthusiastically but the project was delayed for many reasons, one of the most important of which was that I did not have at that time any clear idea as to how the revision was to be carried out. Eventually I decided to leave intact most of the original material, but make the current edition a little more up-to-date by adding, in the form of notes to the individual chapters, some recent references and occasional brief discussions of topics not treated in the original text. The only substantive change from the earlier work is in the treatment of projective geometry; Chapters II through V of the original Volume I have been condensed and streamlined into a single Chapter II. I wish to express my deep gratitude to Donald Babbitt for his generous advice that helped me in organizing this revision, and to Springer-Verlag for their patience and understanding that went beyond what one has a right to expect from a publisher.

I suppose an author's feelings are always mixed when one of his books that is comparatively old is brought out once again. The progress of Science in our time is so explosive that a discovery is hardly made before it becomes obsolete; and yet, precisely because of this, it is essential to keep in sight the origins of things that are taken for granted, if only to lend some perspective to what we are trying to achieve. All I can say is that there are times when one should look back as well as forward, and that the ancient lines, part of which are quoted above still capture the spirit of my thoughts.

V. S. VARADARAJAN

Pacific Palisades,
Dec. 22, 1984

* *Bhagavadgita, 2:47a.*

TABLE OF CONTENTS

Introduction xv

CHAPTER I

Boolean Algebras on a Classical Phase Space 1
1. The Classical Phase Space 1
2. The Logic of a Classical System 6
3. Boolean Algebras 8
4. Functions 12
Notes on Chapter I 17

CHAPTER II

Projective Geometries 18
1. Complemented Modular Lattices 18
2. Isomorphisms of Projective Geometries. Semilinear Transformations 20
3. Dualities and Polarities 22
4. Orthocomplementations and Hilbert Space Structures 26
5. Coordinates in Projective and Generalized Geometries 28
Notes on Chapter II 38

CHAPTER III

The Logic of a Quantum Mechanical System 42
1. Logics 42
2. Observables 45
3. States 48
4. Pure States. Superposition Principle 52
5. Simultaneous Observability 54
6. Functions of Several Observables 62
7. The Center of a Logic 63
8. Automorphisms 67
Notes on Chapter III 70

CHAPTER IV

Logics Associated with Hilbert Spaces 72
1. The Lattice of Subspaces of a Banach Space 72
2. The Standard Logics: Observables and States 80

3. The Standard Logics: Symmetries 104
4. Logics Associated with von Neumann Algebras 112
5. Isomorphism and Imbedding Theorems 114
Notes on Chapter IV 122

CHAPTER V

Measure Theory on G-Spaces 148
1. Borel Spaces and Borel Maps 148
2. Locally Compact Groups. Haar Measure 156
3. G-Spaces 158
4. Transitive G-Spaces 164
5. Cocycles and Cohomology 174
6. Borel Groups and the Weil Topology 191
Notes on Chapter V 200

CHAPTER VI

Systems of Imprimitivity 201
1. Definitions 201
2. Hilbert Spaces of Vector Valued Functions 208
3. From Cocycles to Systems of Imprimitivity 213
4. Projection Valued Measures 217
5. From Systems of Imprimitivity to Cocycles 219
6. Transitive Systems 222
7. Examples and Remarks 228
8. Semidirect Products 236
tes on Chapter VI 241

CHAPTER VII

Multipliers 243
1. The Projective Group 243
2. Multipliers and Projective Representations 247
3. Multipliers and Group Extensions 251
4. Multipliers for Lie Groups 259
5. Examples 275
Notes on Chapter VII 287

CHAPTER VIII

Kinematics and Dynamics 288
1. The Abstract Schrödinger Equation 288
2. Covariance and Commutation Rules 293
3. The Schrödinger Representation 295
4. Affine Configuration Spaces 300
5. Euclidean Systems: Spin 303
6. Particles 312
Notes on Chapter VIII 315

CHAPTER IX

Relativisitic Free Particles 322

1. Relativistic Invariance 322
2. The Lorentz Group 330
3. The Representations of the Inhomogeneous Lorentz Group 343
4. Clifford Algebras 348
5. Representations in Vector Bundles and Wave Equations 356
6. Invariance Under the Inversions 372
7. Localization 377
8. Galilean Relativity 391
Notes on Chapter IX 399

Bibliography 400

Index 407

INTRODUCTION

As laid down by Dirac in his great classic [1], the principle of superposition of states is the fundamental concept on which the quantum theory of atomic systems is to be erected. Dirac's development of quantum mechanics on an axiomatic basis is undoubtedly in keeping with the greatest traditions of the physical sciences. The scope and power of this principle can be recognized at once if one recalls that it survived virtually unmodified throughout the subsequent transition to a relativistic view of the atomic world. It must be pointed out, however, that the precise mathematical nature of the superposition principle was only implicit in the discussions of Dirac; we are indebted to John von Neumann for explicit formulation. In his characteristic way, he discovered that the set of experimental statements of a quantum mechanical system formed a projective geometry—the projective geometry of subspaces of a complex, separable, infinite dimensional Hilbert space. With this as a point of departure, he carried out a mathematical analysis of the axiomatic foundations of quantum mechanics which must certainly rank among his greatest achievements [1] [3] [4] [5] [6].

Once the geometric point of view is accepted, impressive consequences follow. The automorphisms of the geometry describe the dynamical and kinematical structure of quantum mechanical systems, thus leading to the *linear* character of quantum mechanics. The covariance of the physical laws under appropriate space-time groups consequently expresses itself in the form of projective unitary representations of these groups. The economy of thought as well as the unification of method that this point of view brings forth is truly immense; the Schrödinger equation, for example, is obtained from a representation of the time-translation group, the Dirac equation from a representation of the inhomogeneous Lorentz group. This development is the work of many mathematicians and physicists. However, insofar as the mathematical theory is concerned, no contribution is more outstanding than that of Eugene P. Wigner. Beginning with his famous article on time inversion and throughout his great papers on relativistic invariance [1] [3] [4] [5] [6], we find a beautiful and coherent approach to the mathematical description of the quantum mechanical world which achieves nothing less than the fusion of group theory and quantum mechanics, and moreover does this without compromising in any manner the axiomatic principles formulated by Dirac and von Neumann.

My own interest in the mathematical foundations of quantum mechanics received a great stimulus from the inspiring lectures given by Professor George W. Mackey at the University of Washington in Seattle during the summer of 1961. The present volumes are in great part the result of my interest in a detailed elaboration of the main features of the theory sketched by Mackey in those lectures. In sum, my indebtedness to Professor Mackey's lectures and to the books and papers of von Neumann and Wigner is immense and carries through this entire work.

There exist today many expositions of the basic principles of quantum mechanics. At the most sophisticated mathematical level, there are the books of von Neumann [1], Hermann Weyl [1] and Mackey [1]. But, insofar as I am aware, there is no account of the technical features of the geometry and group theory of quantum mechanical systems that is both reasonably self-contained and comprehensive enough to be able to include Lorentz invariance. Moreover, recent re-examinations of the fundamental ideas by numerous mathematicians have produced insights that have substantially added to our understanding of quantum foundations. From among these I want to single out for special mention Gleason's proof that quantum mechanical states are represented by the so-called density matrices, Mackey's extensive work on systems of imprimitivity and group representations, and Bargmann's work on the cohomology of Lie groups, particularly of the physically interesting groups and their extensions. All of this has made possible a conceptually unified and technically cogent development of the theory of quantum mechanical systems from a completely geometric point of view. The present work is an attempt to present such an approach.

Our approach may be described by means of a brief outline of the contents of the three parts that make up this work. The first part begins by introducing the viewpoint of von Neumann according to which every physical system has in its background a certain orthocomplemented lattice whose elements may be identified with the experimentally verifiable propositions about the system. For classical systems this lattice (called the logic of the system) is a Boolean σ-algebra while for quantum systems it is highly nondistributive. This points to the relevance of the theory of complemented lattices to the axiomatic foundations of quantum mechanics. In the presence of modularity and finiteness of rank, these lattices decompose into a direct sum of irreducible ones, called geometries. A typical example of a geometry is the lattice of subspaces of a finite dimensional vector space over a division ring. The theory of these vector geometries is taken up in Chapter II. The isomorphisms of such a geometry are induced in a natural fashion by semilinear transformations. Orthocomplementations are induced by definite semi-bilinear forms which are symmetric with respect to suitable involutive anti-automorphisms of the basic division ring. If the division ring is the reals, complexes or quaternions, this leads to the Hilbert space structures. In this chapter, we also examine the relation between axiomatic geometry and analytic geometry along classical lines with suitable modifications in order to handle the infinite dimensional case also. The main result of this chapter is the theorem which asserts that an abstractly given generalized geometry (i.e., one whose dimension need not be finite) of rank ≥ 4 is isomorphic to the lattice of all finite dimensional subspaces of a vector space over a division ring. The division ring is an invariant of the lattice.

The second part analyzes the structure of the logics of quantum mechanical systems. In Chapter III, we introduce the notion of an abstract logic ($=$ orthocomplemented weakly modular σ-lattice) and the observables and states associated with it. It is possible that certain observables need not be simultaneously observable. It is proved that for a given family of observables to be simultaneously measurable, it is necessary and sufficient that the observables of the family be classically related, i.e., that there exists a Boolean sub σ-algebra of the logic in question to which all the members of the given family are associated. Given an observable and a state, it is shown how to compute the probability distribution of the observable in that state. In Chapter IV, we take up the problem of singling out the logic of all subspaces of a Hilbert space by a set of neat axioms. Using the results of Chapter II, it is proved that the standard logics are precisely the projective ones. The analysis of the notions of an observable and a state carried out in Chapter III now leads to the correspondence between observables and self-adjoint operators, and between the pure states and the rays of the underlying Hilbert space. The automorphisms of the standard logics are shown to be induced by the unitary and antiunitary operators. With this the von Neumann program of a deductive description of the principles of quantum mechanics is completed. The remarkable fact that there is a Hilbert space whose self-adjoint operators represent the observables and whose rays describe the (pure) states is thus finally established to be a consequence of the projective nature of the underlying logic.

The third and final part of the work deals with specialized questions. The main problem is that of a covariant description of a quantum mechanical system, the covariance being with respect to suitable symmetry groups of the system. The theory of such systems leads to sophisticated problems of harmonic analysis on locally compact groups. Chapters V, VI, and VII are devoted to these purely mathematical questions. The results obtained are then applied to yield the basic physical results in Chapters VIII and IX. In Chapter VIII, the Schrödinger equation is obtained and the relations between the Heisenberg and Schrödinger formulations of quantum mechanics are analyzed. The usual expressions for the position, momentum, and energy observables of a quantum mechanical particle are shown to be inevitable consequences of the basic axioms and the requirement of covariance. In addition, a classification of single particle systems is obtained in terms of the spin of the particle. The spin of a particle, which is so characteristic of quantum mechanics, is a manifestation of the *geometry* of the configuration space of the particle.

The final chapter discusses the description of free particles from the relativistic viewpoint. The results of Chapters V, VI and VII are used to obtain a classification of these particles in terms of their mass and spin. With each particle it is possible to associate a vector bundle whose square integrable sections constitute the Hilbert space of the particle. These abstract results lead to the standard transformation formulae for the (one particle) states under the elements of the relativity group. By taking Fourier transforms, it is possible to associate with each particle a definite wave equation. In particular, the Dirac equation of the free electron is obtained in this manner. The same methods lead to the localization in space, for a given time instant, of the particles of nonzero rest mass. The chapter ends with an analysis of Galilean relativity. It is shown that the free particles which are governed by Galilei's principle of relativity are none other than the Schrödinger particles of *positive* mass and arbitrary spin.

With this the program of obtaining a geometric view of the quantum mechanical world is completed. It is my belief that no other approach leads so clearly and smoothly to the fundamental results. It may be hoped that such methods may also lead to a successful description of the world of interacting particles and their fields. The realization of such hopes seems to be a matter for the future.

V. S. VARADARAJAN

CHAPTER I

BOOLEAN ALGEBRAS ON A CLASSICAL PHASE SPACE

1. THE CLASSICAL PHASE SPACE

We begin with a brief account of the usual description of a classical mechanical system with a finite number of *degrees of freedom.* Associated with such a system there is an integer n, and an open set M of the n-dimensional space R^n of n-tuples $(x_1, x_2, \cdots, x_n)$ of real numbers. n is called the *number* of degrees of freedom of the system. The points of M represent the possible *configurations* of the system. A *state* of the system at any instant of time is specified completely by giving a $2n$-tuple $(x_1, x_2, \cdots, x_n, p_1, \cdots, p_n)$ such that $(x_1, \cdots, x_n)$ represents the configuration and $(p_1, \cdots, p_n)$ the momentum vector, of the system at that instant of time. The possible states of the system are thus represented by the points of the open set $M \times R^n$ of R^{2n}. The law of evolution of the system is specified by a smooth function H on $M \times R^n$, called the *Hamiltonian* of the system. If $(x_1(t), \cdots, x_n(t), p_1(t), \cdots, p_n(t))$ represents the state of the system at time t, then the functions $x_i(\cdot)$, $p_i(\cdot)$, $i=1, 2, \cdots, n$, satisfy the well known differential equations:

$$\begin{aligned} \frac{dx_i}{dt} &= \frac{\partial H}{\partial p_i}, \qquad i = 1, 2, \cdots, n, \\ \frac{dp_i}{dt} &= -\frac{\partial H}{\partial x_i}, \qquad i = 1, 2, \cdots, n. \end{aligned} \tag{1}$$

For most of the systems which arise in practice these equations have unique solutions for all t in the sense that given any real number t_0, and a point $(x_1{}^0, x_2{}^0, \cdots, x_n{}^0, p_1{}^0, \cdots, p_n{}^0)$ of $M \times R^n$, there exists a unique differentiable map $t \rightarrow (x_1(t), \cdots, x_n(t), p_1(t), \cdots, p_n(t))$ of R^1 into $M \times R^n$ such that $x_i(\cdot)$ and $p_i(\cdot)$ satisfy the equations (1) with the *initial conditions*

$$x_i(t_0) = x_i{}^0, \quad p_i(t_0) = p_i{}^0, \qquad i = 1, 2, \cdots, n. \tag{2}$$

If we denote by s an arbitrary point of $M \times R^n$, it then follows in the standard fashion that for any t there exists a mapping $D(t)(s \rightarrow D(t)s)$

of $M \times R^n$ into itself with the property that if s is the state of the system at time t_0, $D(t)s$ is the state of the system at time $t+t_0$. The transformations $D(t)$ are one-one, map $M \times R^n$ onto itself and satisfy the equations:

$$\begin{aligned} D(0) &= I \quad \text{(the identity mapping)}, \\ D(-t) &= D(t)^{-1}, \\ D(t_1+t_2) &= D(t_1)D(t_2). \end{aligned} \tag{3}$$

If, in addition, H is an indefinitely differentiable function, then the $D(t)$ are also indefinitely differentiable and the correspondence $t \rightarrow D(t)$ defines a one-parameter differentiable transformation group of $M \times R^n$ so that the map $t, s \rightarrow D(t)s$ of $R^1 \times M \times R^n$ into $M \times R^n$ is indefinitely differentiable. The set $M \times R^n$ of all the possible states of the system is called the *phase space* of the system.

In the formulation described above, the physical quantities or the *observables* of the system are described by real valued functions on $M \times R^n$. For example, if the system is that of a single particle of mass m which moves under some potential field, then $n=3$, $M=R^3$, and the Hamiltonian H is given by

$$H(x_1,x_2,x_3,p_1,p_2,p_3) = \frac{1}{2m}(p_1{}^2+p_2{}^2+p_3{}^2)+V(x_1,x_2,x_3). \tag{4}$$

The function $s \rightarrow (p_1{}^2+p_2{}^2+p_3{}^2)/2m$ is the kinetic energy of the particle and the function $s \rightarrow V(x_1,x_2,x_3)$ is the potential energy of the particle. The function $s \rightarrow p_i$ $(i=1, 2, 3)$ represents the x_i-component of the momentum of the particle. In the general case, if f is a function on $M \times R^n$ which describes an observable, then $f(s)$ gives the value of that observable when the system is in the state s.

This formulation of the basic ideas relating to the mechanics of a classical system can be generalized significantly (Mackey [1], Sternberg [1]). Briefly, this generalization consists in replacing the assumption that M is an open subset of R^n by the more general one that M is an abstract C^∞ manifold of dimension n. The set of all possible configurations of the system is now M, and for any $x \in M$, the momenta of the system at this configuration are the elements of the vector space $M_x{}^*$, which is the dual of the *tangent vector space* M_x of M at x. The phase space of the system is then the set of all possible pairs (x,p), where $x \in M$ and $p \in M_x{}^*$. This set, say S, comes equipped with a natural differentiable structure under which it is a C^∞ manifold of dimension $2n$, the so-called *cotangent bundle* of M. The manifold S admits further a canonical 2-form which is everywhere nonsingular and this gives rise to a natural isomorphism J of the module of all C^∞ vector fields on S onto the module of all 1-forms (both being considered as modules over the ring of C^∞ functions on S). The dynamical development of the system is then specified by a C^∞ function H on S, the *Hamiltonian* of the system. If $t \to s(t)$ is a curve representing a possible evolution of the system, then we have the differential equations:

$$\frac{ds(t)}{dt} = [J^{-1}(dH)](s(t)). \tag{5}$$

Here $ds(t)/dt$ is the tangent vector to S at the point $s(t)$ along the curve $t \to s(t)$ and $J^{-1}(dH)$ is the vector field on S corresponding to the 1-form dH; the right side of the equation (5) being the value of this vector field at the point $s(t)$ of S. In the special case when M is an open set in R^n and $x_1, x_2, \cdots, x_n$ are the global affine coordinates on M, S is canonically identified with $M \times R^n$ and, under this identification, J goes over into the map which transforms the vector field

$$\sum_{i=1}^{n} A_i(\partial/\partial x_i) + \sum_{i=1}^{n} B_i(\partial/\partial p_i)$$

into the 1-form

$$-\sum_{i=1}^{n} B_i\, dx_i + \sum_{i=1}^{n} A_i\, dp_i.$$

The equation (5) then goes over to (1) (cf. Chevalley [1], Helgason [1] for a discussion of the general theory of differentiable manifolds).

In this general setup, the dynamical development of the system is given by the integral curves of the vector field $J^{-1}(dH)$. It is necessary to assume that the integral curves are defined for *all* values of the time parameter t. One can then use the standard theory of vector fields to deduce the existence of a diffeomorphism $D(t)$ of S for each t such that the correspondence $t \to D(t)$ satisfies the conditions (3), and the map $t, s \to D(t)s$ of $R^1 \times S$ into S is C^∞. If the system is at the state s at time t_0, then its state at time $t+t_0$ is $D(t)s$. The physical observables of the system are then represented by real valued functions on S. A special class of Hamiltonian functions, analogous to (4), may be defined in this general framework. Let $x \to \langle . , . \rangle_x$ be a C^∞ Riemannian metric on M, $\langle . , . \rangle_x$ being a positive definite inner product on $M_x \times M_x$. For each $x \in M$, we then have a natural isomorphism $p \to p^*$ of $M_x{}^*$ onto M_x such that $p(u) = \langle u, p^* \rangle_x$ for all $p \in M_x{}^*$ and for all $u \in M_x$. The analogue of (4) is then the Hamiltonian H given by

$$H(x,p) = \langle p^*, p^* \rangle_x + V(x), \tag{6}$$

where V is a C^∞ function on M. The function $x, p \to \langle p^*, p^* \rangle_x$ then represents the kinetic energy of the system in question.

It may be pointed out that one can introduce the concept of the *momenta* of the system in this setup. Let

$$\gamma : t \to \gamma_t \tag{7}$$

be a one-parameter group of *symmetries* of the configuration space M, i.e., $\gamma(t \to \gamma_t)$ is a one-parameter group of C^∞ diffeomorphisms of M onto itself such that the map $t, x \to \gamma_t(x)$ of $R^1 \times M$ into M is C^∞. The infinitesimal generator of γ is a C^∞ vector field, say X_γ, on M; for any $x \in M$ and any real valued C^∞ function f defined around x,

$$(X_\gamma f)(x) = \left\{\frac{d}{dt} f(\gamma_{-t}(x))\right\}_{t=0}.$$

X_γ defines, in a natural fashion, a C^∞ function μ_γ on S. In fact, if $x \in M$ and $p \in M_x{}^*$,

$$\mu_\gamma(x,p) = p(X_\gamma(x))$$

(here $X_\gamma(x)$ denotes the tangent vector to M at x which is the value of X_γ at x). The observable corresponding to the function μ_γ is called the *momentum* of the system *corresponding to the one-parameter group of symmetries* γ. If $M = R^n$, if $x_1, \cdots, x_n$ are the global affine coordinates on M, and if

$$\gamma_t{}^{c}(x_1, \cdots, x_n) = x_1 - tc_1, \cdots, x_n - tc_n,$$

then the observable corresponding to μ_{γ^c} is called the *component of the linear momentum along* $(c_1, \cdots, c_n)$. In the same case, if

$$\gamma_t^{i,j}(x_1, \cdots, x_n) = (y_1, \cdots, y_n),$$

where

$$y_r = x_r, \quad r \neq i, j,$$

$$y_i = x_i \cos t + x_j \sin t, \quad y_j = -x_i \sin t + x_j \cos t,$$

then the observable corresponding to $\mu_{\gamma^{i,j}}$ is called the *angular momentum with respect to a rotation in the i-j plane*. A straightforward calculation shows that in the case when $M = R^n$, $S = R^n \times R^n$, and $x_1, \cdots, x_n, p_1, \cdots, p_n$ are global coordinates on S ($(x_1, \cdots, x_n, p_1, \cdots, p_n)$ depicts $\sum_{i=1}^n p_i(dx_i)_x$),

$$\mu_{\gamma^c}(x,p) = c_1 p_1 + \cdots + c_n p_n,$$

and

$$\mu_{\gamma^{i,j}}(x,p) = x_i p_j - x_j p_i.$$

Suppose now that M is a general C^∞ manifold and S its cotangent bundle. If f and g are two C^∞ functions on S, then we can form $J^{-1}(df)$, which is a C^∞ vector field on S, and apply it to g to get another C^∞ function on S, denoted by $[f,g]$:

$$[f,g] = (J^{-1}(df))g. \tag{8}$$

$[f,g]$ is called the *Poisson Bracket* of f with g. If we use local coordinates $x_1, \cdots, x_n$ on M and the induced coordinates $x_1, \cdots, x_n, p_1, \cdots, p_n$ on S (so that $(x_1, \cdots, x_n, p_1, \cdots, p_n)$ represents $\sum_i p_i(dx_i)$), then J goes over into the map which (locally) sends $\sum_i A_i(\partial/\partial x_i)+\sum_i B_i(\partial/\partial p_i)$ into $-\sum_i B_i\,dx_i+\sum_i A_i\,dp_i$, and $[f,g]$ becomes

$$[f,g] = \sum_i \left(\frac{\partial f}{\partial p_i}\frac{\partial g}{\partial x_i}-\frac{\partial f}{\partial x_i}\frac{\partial g}{\partial p_i}\right). \tag{9}$$

The map $f, g \rightarrow [f,g]$ is bilinear, skew symmetric, and satisfies the identity

$$[[f,g],h]+[[g,h],f]+[[h,f],g] = 0,$$

as is easily verified from (9). If X is any C^∞ vector field on M and μ_X is the C^∞ function on S defined by

$$\mu_X(x,p) = p(X(x)),$$

then one can verify, using (9), that

$$\mu_{aX+bY} = a\mu_X+b\mu_Y \qquad (a,\, b \text{ constants}),$$
$$\mu_{[X,Y]} = [\mu_X,\mu_Y],$$

where $[X,Y]$ is the *Lie bracket* of the vector fields X and Y. If f is any C^∞ function on M and f^0 is the lifted function on S, i.e.,

$$f^0(x,p) = f(x),$$

then we may use (9) once again to check that

$$[\mu_X,f^0] = (Xf)^0$$

for any C^∞ vector field X on M.

In many problems, there is a Lie group G which acts on M and provides the natural symmetries of the problem. For $g \in G$ we write $x \rightarrow g\cdot x$ for the symmetry associated with g and assume that $g, x \rightarrow g\cdot x$ is C^∞ from $G\times M$ into M. In such problems, one restricts oneself to the momenta specified by the one-parameter groups of M. If $\mathfrak{g}$ is the Lie algebra of G (cf. Chevalley [1]) and if we associate for $X \in \mathfrak{g}$, the vector field on M denoted by X also and defined by

$$(Xf)(x) = \left(\frac{d}{dt}f(\exp tX\cdot x)\right)_{t=0},$$

then we obtain the relations

$$\begin{aligned} \mu_{[X,Y]} &= [\mu_X,\mu_Y], \\ [\mu_X,f^0] &= (Xf)^0 \end{aligned} \tag{10}$$

between the configuration observables f^0 and the momentum observables μ_X. These relations are usually referred to as *commutation rules*.

2. THE LOGIC OF A CLASSICAL SYSTEM

We shall now examine the algebraic aspects of a general classical system. In view of the discussion carried out just now, it is clear that for any classical system $\mathfrak{S}$ there is associated a space S called the phase space of $\mathfrak{S}$. The states of the system are in one-one correspondence with the points of S. The notion of a state is so formulated that if one knows the state of the system at an instant of time t_0, and also the dynamical law of evolution of the system, then one can determine the state of the system at time $t+t_0$. The observables or physical quantities which are of interest to the observer are then represented by real valued functions on S. If f is the function corresponding to a particular observable, its value $f(s)$ at the point s of S is interpreted as the value of the physical quantity when the system is in the state s. If s is the state of the system at time t_0, we can write $D(t)s$ for the state of the system at time $t+t_0$. We thus have a transformation $D(t)$ of S into itself. For each t, $D(t)$ is invertible and maps S onto itself. The correspondence $t \rightarrow D(t)$ satisfies the equations (3). $t \rightarrow D(t)$ is then a one-parameter group of transformations of S. It is called the *dynamical group* of the system $\mathfrak{S}$.

These concepts make sense in every classical system. In the case of any such system the most general statement which can be made about it is one which asserts that the value of a certain observable lies in a real number set E. If the observable is represented by the function f on S, then such a statement is equivalent to the statement that the state of the system lies in the set $f^{-1}(E)$ of the space S. In other words, the physically meaningful statements that can be made about the system are in correspondence with certain subsets of S. The inclusion relations for subsets naturally correspond to implications of statements. In mathematical terms, this means that at the background of the classical system there is a Boolean algebra of subsets of the space S, the elements of which represent the statements about the physical system. It is natural to call this Boolean algebra the *logic* of the system.

Suppose now that $\mathfrak{S}$ is a system which does not follow the laws of classical mechanics. Then one cannot associate with it a phase space in general. It is nevertheless meaningful to consider the totality of experimentally verifiable statements which may be made about the system. This collection, which may be called the *logic* of $\mathfrak{S}$, comes equipped with the relations of implication and negation which convert it into a complemented partially ordered set. For a classical system this partially ordered set is a Boolean algebra. Clearly, it is possible to conceive of mechanical systems whose logics are not Boolean algebras. *We take the point of view that quantum mechanical systems are those whose logics form some sort of projective geometries and which are consequently nondistributive* *lattices*. With such a point of view it is possible to understand the role played by simultaneously observable quantities, the uncertainty relations, and the complementarity principles. These phenomena, which are so peculiar to quantum systems, will then be seen to be consequences of the nondistributive nature of the logic in the background of the system $\mathfrak{S}$.

It might seem a bit surprising that the basic assumption on a quantum system is that its logic is not a distributive lattice. It would be natural to argue that statements about a physical system should obey the same rules as the rules of ordinary set theory. The well known critiques of von Neumann and Heisenberg address this question (von Neumann [1], Birkhoff-von Neumann [1], Heisenberg [1]). The point is that only *experimentally verifiable statements* are to be regarded as members of the logic of the system. Consequently, as it happens in many questions in atomic physics, it may be impossible to verify *experimentally* statements which involve the values of two physical quantities of the system—for example, measurements of the position and momentum of an electron. One can verify statements about one of them but not, in general, those which involve both of them. What the basic assumptions imply is that the statements regarding position or momentum form two Boolean subalgebras of the logic but that there is in general no Boolean algebra which contains both of these Boolean subalgebras.

Before beginning an analysis of the logic of general quantum mechanical systems it would be helpful to recast at least some of the features of the formulation given in section 1 in terms of the logic of the classical system. In the first place it is natural to strengthen the hypothesis and assume that the logic of a given classical system $\mathfrak{S}$ is a Boolean σ-algebra, say $\mathscr{L}$, of subsets of S, the phase space of $\mathfrak{S}$. Suppose now, that an observable associated with the system is represented by the real valued function f on S. The statements concerning the observable are then those which assert that its value lies in an arbitrary Borel set E of the real line and these are represented by the subsets $f^{-1}(E)$ of S. The observable can thus be represented, without any loss of physical content, equally by the map $E \rightarrow f^{-1}(E)$ of the class of Borel subsets of the real line into $\mathscr{L}$. The range of this mapping is a sub-σ-algebra, say $\mathscr{L}_f$. Suppose g is a real valued Borel function on the real line. Then, the observable represented by the function $g \circ f$ $(s \rightarrow g(f(s)))$ can also be represented by the map $E \rightarrow f^{-1}(g^{-1}(E))$ from which we conclude that $\mathscr{L}_{g \circ f}$ is contained in $\mathscr{L}_f$.

In order to formulate the general features of a classical mechanical system in terms of its logic $\mathscr{L}$, it is therefore necessary to determine to what extent an abstract σ-algebra $\mathscr{L}$ can be regarded as a σ-algebra of subsets of some space S; further to determine the class of mappings from the σ-algebra of Borel sets of the real line into $\mathscr{L}$ which correspond to real valued functions on S; and to clarify the concept of functional dependence in this general context. We shall now proceed to a discussion of these questions.

3. BOOLEAN ALGEBRAS

Let $\mathscr{L}$ be a nonempty set. $\mathscr{L}$ is said to be *partially ordered* if there is a relation $<$ between some pairs of elements of $\mathscr{L}$ such that (i) $a<a$ for all a in $\mathscr{L}$; (ii) $a<b$ and $b<a$ imply $a=b$; (iii) $a<b$ and $b<c$ imply $a<c$. If $\mathscr{L}$ is partially ordered, there is at most one element called the *null* or *zero* element and denoted by 0, such that $0<a$ for all a in $\mathscr{L}$. Similarly there is at most one element called the *unit* element and denoted by 1, such that $a<1$ for all a in $\mathscr{L}$. More generally, for any nonempty subset F of $\mathscr{L}$ there exists at most one element c of $\mathscr{L}$ such that (i) $a<c$ for all $a \in F$; (ii) if d is any element of $\mathscr{L}$ such that $a<d$ for all $a \in F$, then $c<d$. We shall write $\bigvee_{a\in F} a$ for c whenever it exists. If F is a finite set, say $F=\{a_1, \cdots, a_n\}$, it is customary to write $\bigvee_{i=1}^n a_i$ or $a_1 \vee a_2 \vee \cdots \vee a_n$ instead of $\bigvee_{a\in F} a$. In an analogous fashion, for any subset F of $\mathscr{L}$ there exists at most one element c such that (i) $c<a$ for all $a \in F$; (ii) if d is any element of $\mathscr{L}$ such that $d<a$ for all $a \in F$, then $d<c$; we denote it by $\bigwedge_{a\in F} a$ whenever it exists. If F is a finite set, say $F=\{a_1, \cdots, a_n\}$, we often write $\bigwedge_{i=1}^n a_i$ or $a_1 \wedge a_2 \wedge \cdots \wedge a_n$ instead of $\bigwedge_{a\in F} a$. The partially ordered set $\mathscr{L}$ is called a *lattice* if

$$(11) \qquad \begin{aligned} &\text{(i)} \quad 0 \text{ and } 1 \text{ exist in } \mathscr{L} \text{ and } 0 \neq 1, \\ &\text{(ii)} \quad \bigvee_{a\in F} a \text{ and } \bigwedge_{a\in F} a \text{ exist for all finite subsets } F \text{ of } \mathscr{L}. \end{aligned}$$

Suppose that $\mathscr{L}$ is a lattice. Given any element a of $\mathscr{L}$, an element a' of $\mathscr{L}$ is said to be a *complement* of a if $a \wedge a'=0$ and $a \vee a'=1$. a is then a complement of a'. $\mathscr{L}$ is said to be *complemented* if, given any element, there exists at least one complement of it. It is obvious that 0 and 1 have the unique complements 1 and 0, respectively. A lattice $\mathscr{L}$ is said to be *distributive* if for any three elements a, b, c of $\mathscr{L}$, the identities

$$a \wedge (b \vee c) = (a \wedge b) \vee (a \wedge c),$$
$$a \vee (b \wedge c) = (a \vee b) \wedge (a \vee c)$$

are satisfied. A complemented distributive lattice is called a *Boolean algebra*. A *Boolean σ-algebra* $\mathscr{L}$ is a Boolean algebra in which $\bigwedge_{a\in F} a$ and $\bigvee_{a\in F} a$ exist for every countable subset F of $\mathscr{L}$.

Every element in a Boolean algebra has a unique complement. Suppose in fact that $\mathscr{L}$ is a Boolean algebra and that a is an element with two complements a_1 and a_2. Then, one has

$$a_1 = a_1 \wedge (a \vee a_2) = (a_1 \wedge a) \vee (a_1 \wedge a_2) = a_1 \wedge a_2 < a_2;$$

similarly, $a_2 < a_1$, so that $a_1 = a_2$. The unique complement of a is denoted by a'. Using the standard manipulations of set theory it is easy to show that $(\bigwedge_{a \in F} a)' = \bigvee_{a \in F} a'$ and $(\bigvee_{a \in F} a)' = \bigwedge_{a \in F} a'$ for any finite subset F of $\mathscr{L}$. If $\mathscr{L}$ is a Boolean σ-algebra, then the same identities remain valid even when F is countably infinite. If $\mathscr{L}$ is any Boolean algebra and a, b are elements in it with $a < b$, $c = a' \wedge b$ is the unique element of $\mathscr{L}$ such that $a \wedge c = 0$ and $a \vee c = b$; c is called *the complement of a in b*. Since $c = b \wedge a'$, $c < a'$ (cf. Birkhoff [1] and Sikorski [1] for the general theory of Boolean algebras and σ-algebras).

A *homomorphism* of a Boolean algebra $\mathscr{L}_1$ into a Boolean algebra $\mathscr{L}_2$ is a map h of $\mathscr{L}_1$ into $\mathscr{L}_2$ such that (i) $h(0) = 0$, $h(1) = 1$; (ii) $h(a') = h(a)'$ for all a in $\mathscr{L}_1$; (iii) $h(a \vee b) = h(a) \vee h(b)$, $h(a \wedge b) = h(a) \wedge h(b)$ for all a, b in $\mathscr{L}_1$. If h is a homomorphism and $a < b$, then $h(a) < h(b)$. An *isomorphism* of $\mathscr{L}_1$ onto $\mathscr{L}_2$ is a homomorphism h of $\mathscr{L}_1$ onto $\mathscr{L}_2$ such that $h(a) = 0$ if and only if $a = 0$; in this case h is also one-one.

The class of *all* subsets of any set is a Boolean algebra under set inclusion and set complementation. However, obviously this is not the most general Boolean algebra since infinite unions and intersections exist in it. Suppose now that X is a topological space. The class of subsets of X which are both open and closed (open-closed) is obviously a Boolean algebra. A well known theorem of Stone [1] asserts that every Boolean algebra is isomorphic to one such and that, if we require the topological space to be compact Hausdorff as well as totally disconnected, it is essentially uniquely determined by the Boolean algebra. We recall that a compact space is said to be *totally disconnected* if every open subset of it can be written as a union of open-closed subsets. We shall call a compact Hausdorff totally disconnected space a *Stone space.*

Let $\mathscr{L}$ be a Boolean algebra. A subset $\mathscr{M}$ of $\mathscr{L}$ is called a *dual ideal* if the following properties are satisfied:

(i) $0 \notin \mathscr{M}$,
(ii) if $a \in \mathscr{M}$ and $a < b$, then $b \in \mathscr{M}$,
(iii) if $a, b \in \mathscr{M}$, then $a \wedge b \in \mathscr{M}$.

$\mathscr{M}$ is said to be *maximal* if it is properly contained in no other dual ideal.

The naturalness of the notion of maximal dual ideals can be seen in the following way. Let X be a Stone space and $\mathscr{L} = \mathscr{L}(X)$ the Boolean algebra of all open-closed subsets of X. Then, for any $x \in X$, the collection $\mathscr{M}(x)$, where

$$\mathscr{M}(x) = \{A : A \in \mathscr{L}, x \in A\},$$

is easily seen to be a maximal dual ideal; it is also easy to check that the correspondence $x \rightarrow \mathscr{M}(x)$ is one-one if we notice that X is Hausdorff. The concept of maximal dual ideals is central in the proof of Stone's theorem.

Suppose that $\mathscr{L}$ is an arbitrary Boolean algebra. Using Zorn's lemma one can show easily that maximal dual ideals of $\mathscr{L}$ exist. Let $X = X(\mathscr{L})$ be the set of all maximal dual ideals of $\mathscr{L}$. For any $a \in \mathscr{L}$ we define X_a by

$$X_a = \{\mathscr{M} : \mathscr{M} \in X, a \in \mathscr{M}\},$$

$X_0 = \varnothing$, the null set, and $X_1 = X$. We shall say that a subset $A \subseteq X$ is *open* if A is the union of sets of the form X_a. This definition defines the structure of a topology on X called the *Stone topology*. We now have:

Theorem 1.1 (*Stone* [1]). *Let $\mathscr{L}$ be a Boolean algebra and let $X = X(\mathscr{L})$ be the space of all maximal dual ideals of $\mathscr{L}$. Then, equipped with the Stone topology, X becomes a Stone space. The map $a \rightarrow X_a$ is then an isomorphism of $\mathscr{L}$ with the Boolean algebra of all open-closed subsets of X. X is determined by $\mathscr{L}$, among the class of Stone spaces, up to a homeomorphism. More generally, let X and Y be Stone spaces and let $\mathscr{L}(X)$ and $\mathscr{L}(Y)$ be their respective Boolean algebras of open-closed subsets. If u is any isomorphism of $\mathscr{L}(Y)$ onto $\mathscr{L}(X)$, there exists a homeomorphism h of X onto Y such that*

$$u(A) = h^{-1}(A) \qquad (A \in \mathscr{L}(Y)); \tag{12}$$

moreover, h is uniquely determined by (12).

This theorem is very well known and we do not give its proof. The reader may consult the books of Birkhoff [1], Sikorski [1], and the paper of Stone [1] for the proof.

Corollary 1.2. *Let X be a Stone space and let $\mathscr{L} = \mathscr{L}(X)$ be the Boolean algebra of open-closed subsets of X. If $t \rightarrow D_t(-\infty < t < \infty)$ is any one-parameter group of automorphisms of $\mathscr{L}$, there exists a unique one-parameter group $t \rightarrow h_t$ of homeomorphisms of X onto itself such that for all t and $A \in \mathscr{L}$, $D_t(A) = h_t^{-1}(A)$.*

Proof. Theorem 1.1 ensures the existence and uniqueness of each h_t. If t_1, t_2 are real, then $h_{t_1+t_2}$ and $h_{t_1} \circ h_{t_2}$ induce the same automorphism $D_{t_1+t_2}$ of $\mathscr{L}$, so that $h_{t_1+t_2} = h_{t_1} \circ h_{t_2}$.

The theorem of Stone shows that there is essentially no distinction between an abstract Boolean algebra and a Boolean algebra of sets. If one deals with Boolean σ-algebras, the situation becomes somewhat less straightforward. We shall now describe the modifications necessary when one replaces Boolean algebras by Boolean σ-algebras.

If $\mathscr{L}_1$ and $\mathscr{L}_2$ are Boolean σ-algebras, and h a map of $\mathscr{L}_1$ into $\mathscr{L}_2$, h is called a *σ-homomorphism* if (i) $h(0) = 0$, $h(1) = 1$; (ii) $h(a') = h(a)'$ for all $a \in \mathscr{L}_1$; and (iii) if F is any subset of $\mathscr{L}_1$ which is finite or countably infinite, $h(\bigvee_{a\in F} a) = \bigvee_{a \in F} h(a)$ and $h(\bigwedge_{a\in F} a) = \bigwedge_{a\in F} h(a)$. Suppose $\mathscr{L}_1$

and $\mathscr{L}_2$ are two Boolean σ-algebras and h a σ-homomorphism of $\mathscr{L}_1$ onto $\mathscr{L}_2$. The set $\mathscr{N}=\{a : a \in \mathscr{L}_1, h(a)=0\}$ is a subset of $\mathscr{L}_1$ with the properties: (i) $0 \in \mathscr{N}$, $1 \notin \mathscr{N}$; (ii) if $a \in \mathscr{N}$ and $b<a$, then $b \in \mathscr{N}$; (iii) if F is a countable subset of $\mathscr{N}$, $\bigvee_{a\in F} a \in \mathscr{N}$. $\mathscr{N}$ is called the *kernel* of h. Suppose conversely $\mathscr{L}$ is a Boolean σ-algebra and $\mathscr{N}$ a subset of $\mathscr{L}$ with properties (i) to (iii) listed above. We shall say that elements a and b of $\mathscr{L}$ are equivalent, $a \sim b$, if $a \wedge b'$ and $b \wedge a'$ are in $\mathscr{N}$. It is easily verified that $\sim$ is an equivalence relation. Let $\bar{\mathscr{L}}$ be the set of all equivalence classes, and for any a in $\mathscr{L}$, let $\bar{a}$ denote the unique equivalence class containing a. We define $\bar{a}<\bar{b}$ whenever there are elements a in $\bar{a}$, and b in $\bar{b}$ such that $a<b$. It is then easily shown that $\bar{\mathscr{L}}$ is a Boolean σ-algebra whose zero and unit elements are, respectively, $\bar{0}$ and $\bar{1}$, and that the map $a \rightarrow \bar{a}$ is a σ-homomorphism of $\mathscr{L}$ onto $\bar{\mathscr{L}}$ with kernel $\mathscr{N}$. We write $\bar{\mathscr{L}}=\mathscr{L}/\mathscr{N}$.

Theorem 1.3 (*Loomis* [1]). *Let $\mathscr{L}$ be a Boolean σ-algebra. Then there exists a set X, a σ-algebra $\mathscr{S}$ of subsets of X, and a σ-homomorphism h of $\mathscr{S}$ onto $\mathscr{L}$.*

Proof. Let X be a Stone space such that the lattice $\mathscr{L}'=\mathscr{L}(X)$ of open-closed subsets of X is isomorphic to the Boolean algebra $\mathscr{L}$. Let $\mathscr{S}$ denote the smallest σ-algebra of subsets of X containing $\mathscr{L}'$. We denote by $\cup$ and $\cap$ the operations of set union and set intersection for subsets of X, and by $\vee$ and $\wedge$ the lattice-theoretic operations in $\mathscr{L}$ and $\mathscr{L}'$.

If $A_1, A_2, \cdots$ is any sequence of sets in $\mathscr{L}'$, then $\bigvee_n A_n=A$ exists in $\mathscr{L}'$ since $\mathscr{L}'$ is isomorphic to $\mathscr{L}$ and $\mathscr{L}$ is a σ-algebra. Since A is the smallest element of $\mathscr{L}'$ containing all the A_n, it follows that the set $A-\bigcup_n A_n$ cannot contain any element of $\mathscr{L}'$ as a subset. The sets in $\mathscr{L}'$ form a base for the topology of X and hence we conclude that $A-\bigcup_n A_n$ cannot contain any nonnull open set. Since $\bigcup_n A_n$ is open, this shows that $A-\bigcup_n A_n$ is a closed nondense set.

Consider now the class $\mathscr{S}_1$ of all sets $A \in \mathscr{S}$ with the property that for some B in $\mathscr{L}'$, $(A-B) \cup (B-A)$ is of the first category. If B_1 and B_2 are elements of $\mathscr{L}'$ such that $(A-B_i) \cup (B_i-A)$ is of the first category $(i=1, 2)$, then it will follow that $(B_1-B_2) \cup (B_2-B_1)$ is of the first category, which is not possible (by the category theorem of Baire) unless $B_1=B_2$. Thus, for any A in $\mathscr{S}_1$ there exists a *unique* $B=h_1(A)$ in $\mathscr{L}'$ such that $(A-B) \cup (B-A)$ is of the first category. Clearly $\mathscr{L}' \subseteq \mathscr{S}_1$, and for $A \in \mathscr{L}'$, $h_1(A)=A$.

We claim that $\mathscr{S}_1$ is a σ-algebra. Since

$$(A-B) \cup (B-A) = (A'-B') \cup (B'-A')$$

(primes denoting complementation in X), we see that for any A in $\mathscr{S}_1$,

A' is in $\mathscr{S}_1$ and $h_1(A')=h_1(A)'$. Suppose $A_1, A_2, \cdots$ is any sequence in $\mathscr{S}_1$. Write $B_n=h_1(A_n)$, $A=\bigcup_n A_n$, $B=\bigvee_n B_n$, $B_0=\bigcup_n B_n$. By what we said above, $B-B_0$ is a closed nondense set. Moreover, as $B_0 \subseteq B$, we have

$$\begin{aligned}(A-B) \cup (B-A) &\subseteq \{(A-B_0) \cup (B_0-A)\} \cup (B-B_0)\\ &\subseteq \bigcup_n \{(A_n-B_n) \cup (B_n-A_n)\} \cup (B-B_0).\end{aligned}$$

As all members of the right side are of the first category, this proves that $A \in \mathscr{S}_1$ and $h_1(A)=\bigvee_n h_1(A_n)$. In a similar fashion we can show that $\bigcap_n A_n$ lies in $\mathscr{S}_1$ and $h_1(\bigcap_n A_n)=\bigwedge_n h_1(A_n)$.

The conclusions of the preceding paragraph show that $\mathscr{S}_1$, is a Boolean σ-algebra $\subseteq \mathscr{S}$. Since $\mathscr{S}_1$ contains $\mathscr{L}'$, $\mathscr{S}_1=\mathscr{S}$. Moreover, we see at the same time that h_1 is a σ-homomorphism of $\mathscr{S}$ onto $\mathscr{L}'$. If we write $h=k \circ h_1$ where k is an isomorphism of $\mathscr{L}'$ onto $\mathscr{L}$, then h is a σ-homomorphism of $\mathscr{S}$ onto $\mathscr{L}$.

Remark. Let $\mathscr{S}$ be the σ-algebra of Borel sets on the unit interval [0,1], $\mathscr{N}$ the class of Borel sets of Lebesgue measure 0, and $\mathscr{L}=\mathscr{S}/\mathscr{N}$. Then $\mathscr{L}$ is a Boolean σ-algebra. We can obviously define Lebesgue measure λ as a countably additive function $\bar{\lambda}$ on $\mathscr{L}$; $\bar{\lambda}$ is strictly positive in the sense that for any $a \neq 0$ of $\mathscr{L}$, $\bar{\lambda}(a)$ is positive. From this it follows that any family of mutually disjoint elements of $\mathscr{L}$ is countable. On the other hand, since $\mathscr{S}$ is countably generated, so is $\mathscr{L}$. However, any σ-algebra of subsets of some space X which is countably generated can be proved to have atoms, that is, minimal elements. Since $\mathscr{L}$ does not have atoms, $\mathscr{L}$ cannot be isomorphic to any σ-algebra of sets.

4. FUNCTIONS

We now take up the second question raised in section 2, namely, the problem of describing the calculus of functions on a set X entirely in terms of the Boolean σ-algebra of subsets of X with respect to which all these functions are measurable. The results are summarized in theorems 1.4 and 1.6 of this section.

Let X be any set of points x and $\mathscr{S}$ a Boolean σ-algebra of subsets of X. A function f from X into a complete separable metric space Y is said to be $\mathscr{S}$-measurable if $f^{-1}(E) \in \mathscr{S}$ for all Borel sets $E \subseteq Y$. If f is $\mathscr{S}$-measurable, the mapping $E \rightarrow f^{-1}(E)$ is a σ-homomorphism of the σ-algebra $\mathscr{B}(Y)$ of Borel subsets of Y into $\mathscr{S}$. Suppose now $\mathscr{L}$ is an abstract Boolean σ-algebra. We shall then define a *Y-valued observable associated with $\mathscr{L}$* to be any σ-homomorphism of $\mathscr{B}(Y)$ into $\mathscr{L}$. If $Y=R^1$, the real line, we call these observables *real valued* and refer to them simply as *observables*. From our definition of σ-homomorphisms we see that a map $u(E \rightarrow u(E))$

of $\mathscr{B}(Y)$ into $\mathscr{L}$ is a Y-valued observable associated with $\mathscr{L}$ if and only if (i) $u(\varnothing)=0$, $u(Y)=1$; (ii) $u(Y-E)=u(E)'$ for all E in $\mathscr{B}(Y)$; (iii) if $E_1, E_2, \cdots$ is any sequence of Borel sets in Y, $u(\bigcup_n E_n)=\bigvee_n u(E_n)$ and $u(\bigcap_n E_n)=\bigwedge_n u(E_n)$.

Theorem 1.4 *Let X be a set, $\mathscr{S}$ a Boolean σ-algebra of subsets of X and h a σ-homomorphism of $\mathscr{S}$ onto a Boolean σ-algebra $\mathscr{L}$. Suppose further that $u(E \to u(E))$ is any (real valued) observable associated with $\mathscr{L}$. Then there exists an $\mathscr{S}$-measurable real valued function f defined on X such that*

$$u(E) = h(f^{-1}(E)) \tag{13}$$

for all Borel sets $E \subseteq R^1$. f is essentially unique in the sense that if g is any $\mathscr{S}$-measurable real valued function defined on X such that $u(E)=h(g^{-1}(E))$ for all Borel sets $E \subseteq R^1$, the set $\{x : x \in X, f(x) \neq g(x)\}$ belongs to the kernel of h.

Proof. We begin with a simple observation. Suppose A and B are two subsets of X in $\mathscr{S}$ such that $A \subseteq B$, and c any element of $\mathscr{L}$ such that $h(A) < c < h(B)$. Then we can select a set C in $\mathscr{S}$ such that $A \subseteq C \subseteq B$ and $h(C)=c$. In fact, since h maps $\mathscr{S}$ onto $\mathscr{L}$, there exists C_1 in $\mathscr{S}$ such that $h(C_1)=c$. If we define $C=(C_1 \cap B) \cup A$, then $A \subseteq C \subseteq B$ while

$$h(C) = (h(C_1) \wedge h(B)) \vee h(A) = (c \wedge b) \vee a = c.$$

We now come to the proof of theorem 1.4. Let $r_1, r_2, \cdots$ be any distinct enumeration of the rational numbers in R^1 and let D_i be the interval $\{t : t \in R^1, t < r_i\}$. Evidently, $u(D_i) < u(D_j)$ whenever $r_i < r_j$. We shall now construct sets $A_1, A_2, \cdots$ in $\mathscr{S}$ such that (a) $h(A_i)=u(D_i)$ for all $i=1, 2, 3, \cdots$; (b) $A_i \subseteq A_j$ whenever $r_i < r_j$. Let A_1 be any set in $\mathscr{S}$ such that $h(A_1)=u(D_1)$. Suppose $A_1, A_2, A_3, \cdots, A_n$ in $\mathscr{S}$ have been constructed such that (i) $h(A_i)=u(D_i)$ for $i=1, 2, \cdots, n$; (ii) $A_i \subseteq A_j$ whenever $r_i < r_j$, $1 \leq i, j \leq n$. We shall construct A_{n+1} as follows. Let $(i_1, i_2, \cdots, i_n)$ be the permutation of $(1, 2, \cdots, n)$ such that $r_{i_1} < r_{i_2} < \cdots < r_{i_n}$. Then, there exists a unique k such that $r_{i_k} < r_{n+1} < r_{i_{k+1}}$ (we define $r_{i_0} = -\infty$ and $r_{i_{n+1}} = +\infty$), and by the observation made in the preceding paragraph, we can select A_{n+1} in $\mathscr{S}$ such that $A_{i_k} \subseteq A_{n+1} \subseteq A_{i_{k+1}}$ (we define $A_{i_0} = \varnothing$, $A_{i_{n+1}} = X$). The collection $\{A_1, A_2, \cdots, A_{n+1}\}$ then has the same properties relative to $r_1, r_2, \cdots, r_{n+1}$ as $\{A_1, \cdots, A_n\}$ had relative to $r_1, r_2, \cdots, r_n$. It thus follows by induction that there exists a sequence $A_1, A_2 \cdots$ of sets in $\mathscr{S}$ with the properties (a) and (b). As

$$h\Big(\bigcap_j A_j\Big) = \bigwedge_j u(D_j) = u\Big(\bigwedge_j D_j\Big) = 0,$$

we may, by replacing A_k by $A_k - \bigcap_j A_j$ if necessary, assume that

$\bigcap_j A_j = \varnothing$. Further $h(\bigcup_j A_j) = \bigvee_j u(D_j) = u(\bigvee_j D_j) = 1$ so that $h(N) = 0$, where $N = X - \bigcup_j A_j$. We now define a function f as follows:

$$f(x) = \begin{cases} 0 & \text{if } \quad x \in N, \\ \inf\{r_j : x \in A_j\} & \text{if } \quad x \in \bigcup_j A_j. \end{cases} \tag{14}$$

Clearly, f is finite everywhere. Moreover, for any k,

$$f^{-1}(D_k) \cap (X - N) = \bigcup_{j : r_j < r_k} A_j,$$

so that f is $\mathscr{S}$-measurable. Further,

$$\begin{aligned} h(f^{-1}(D_k)) &= h\Big(\bigcup_{j : r_j < r_k} A_j\Big) \\ &= \bigvee_{j : r_j < r_k} u(D_j) \\ &= u(D_k), \end{aligned}$$

so that $h(f^{-1}(E)) = u(E)$ whenever $E = D_k$ for some k. Since the class of all E for which this equation is valid is a Boolean σ-algebra, we conclude that $h(f^{-1}(E)) = u(E)$ for all Borel sets E.

It remains to examine the uniqueness. Let g be any real valued $\mathscr{S}$-measurable function on X such that $h(g^{-1}(E)) = u(E)$ for all Borel sets E. Then, if we write D_k' for $R^1 - D_k$,

$$\begin{aligned} M &= \{x : x \in X, f(x) \neq g(x)\} \\ &= \bigcup_k \{(f^{-1}(D_k) \cap g^{-1}(D_k')) \bigcup (f^{-1}(D_k') \cap g^{-1}(D_k))\}, \end{aligned}$$

so that

$$\begin{aligned} h(M) &= \bigvee_k \{u(D_k) \wedge u(D_k')\} \\ &= 0. \end{aligned}$$

This shows that M belongs to the kernel of h. This completes the proof of the theorem.

Lemma 1.5. *Let X be a set, $\mathscr{S}$ a σ-algebra of subsets of X and f an $\mathscr{S}$-measurable mapping of X into R^n. Suppose $\mathscr{S}^\sim = \{f^{-1}(F) : F \in \mathscr{B}(R^n)\}$. Then to any $\mathscr{S}^\sim$-measurable real function c on X there corresponds a real valued Borel function $c^\sim$ on R^n such that $c(x) = c^\sim(f(x))$ for all $x \in X$.*

Proof. Since c is $\mathscr{S}^\sim$-measurable, there exists a sequence c_n $(n = 1, 2, \cdots)$ of $\mathscr{S}^\sim$-measurable functions such that (i) each c_n takes only finitely many values; (ii) $c_n(x) \to c(x)$ for all $x \in X$. For any n, let $A_{n1}, A_{n2}, \cdots, A_{nk_n}$ be disjoint subsets of X whose union is X such that c_n is a constant, say a_{ni}, on A_{ni}, the a_{ni} being distinct for $i = 1, 2, \cdots, k_n$. Since c_n is $\mathscr{S}^\sim$-measurable, $A_{ni} \in \mathscr{S}^\sim$; so there exists a Borel set B_{ni} of R^n such that $A_{ni} = f^{-1}(B_{ni})$

$(i=1, 2, \cdots, k_n)$. Replacing B_{ni} by $B_{ni}-\bigcup_{j<i} B_{nj}$ if necessary, we may assume that the B_{ni} are disjoint. Let us define the function $c_n^{\sim}$ on R^n as follows:

$$c_n^{\sim}(t) = \begin{cases} a_{ni} & \text{if} \quad t \in B_{ni}, \\ 0 & \text{if} \quad t \notin \bigcup_i B_{ni}. \end{cases}$$

Clearly $c_n^{\sim}$ is Borel and $c_n(x)=c_n^{\sim}(f(x))$ for all x in X. Let us define $c^{\sim}$ on R^n as follows:

$$c^{\sim}(t) = \begin{cases} \lim\limits_{n\to\infty} c_n^{\sim}(t) \text{ if the limit exists,} \\ 0 \text{ otherwise.} \end{cases}$$

Clearly $c^{\sim}$ is Borel. Since $c_n(x)=c_n^{\sim}(f(x))$ and $\lim c_n(x)$ exists and is equal to $c(x)$ for all x in X, $c^{\sim}(f(x))=c(x)$ for all x in X.

Let $\mathscr{L}$ be a Boolean σ-algebra. $\mathscr{L}_1 \subseteq \mathscr{L}$ is said to be a *sub-σ-algebra* if (i) $0, 1 \in \mathscr{L}_1$; (ii) if $a \in \mathscr{L}_1$, then $a' \in \mathscr{L}_1$; (iii) if $a_1, a_2, \cdots$ are in $\mathscr{L}_1$, then $\bigvee_n a_n$ and $\bigwedge_n a_n$ are in $\mathscr{L}_1$. A sub-σ-algebra $\mathscr{L}_1$ is said to be *separable* if there exists a countable subset D of $\mathscr{L}$ such that $\mathscr{L}_1$ is the smallest sub-σ-algebra of $\mathscr{L}$ containing D.

Theorem 1.6. (i) *Let $\mathscr{L}$ be a Boolean σ-algebra and $u(E \to u(E))$ an observable associated with $\mathscr{L}$. Then the range $\mathscr{L}_u=\{u(E) : E \in \mathscr{B}(R^1)\}$ of u is a separable Boolean sub-σ-algebra of $\mathscr{L}$. Conversely, if $\mathscr{L}_1$ is a separable Boolean sub-σ-algebra of $\mathscr{L}$, there exists an observable u associated with $\mathscr{L}$ such that $\mathscr{L}_1$ is the range of u.*

(ii) *Let u_i $(i=1, 2, \cdots, n)$ be observables associated with $\mathscr{L}$, and $\mathscr{L}_i$ $(i=1, 2, \cdots, n)$ their respective ranges. Suppose $\mathscr{L}_0$ is the smallest sub-σ-algebra of $\mathscr{L}$ containing all the $\mathscr{L}_i$. Then there exists a unique σ-homomorphism u of $\mathscr{B}(R^n)$ (the σ-algebra of Borel subsets of the n-dimensional space R^n) onto $\mathscr{L}_0$ such that for any Borel set E of R^1, $u_i(E)=u(p_i^{-1}(E))$, where p_i is the projection $(t_1, t_2, \cdots, t_n) \to t_i$ of R^n onto R^1. If φ is any real valued Borel function on R^n, the map $E \to u(\varphi^{-1}(E))(E \in \mathscr{B}(R^1))$ is an observable associated with $\mathscr{L}$ whose range is contained in $\mathscr{L}_0$. Conversely, if $v(E \to v(E))$ is any observable associated with $\mathscr{L}$ such that the range of v is contained in $\mathscr{L}_0$, there exists a real valued Borel function φ on R^n such that $v(E)=u(\varphi^{-1}(E))$ for all E.*

Proof. If u is an observable with range $\mathscr{L}_u$, $\mathscr{L}_u$ is obviously the smallest sub-σ-algebra of $\mathscr{L}$ containing all the $u(E)$, where E is any open interval of R^1 with rational end points. This shows that $\mathscr{L}_u$ is separable.

Suppose conversely that $\mathscr{L}_1 \subseteq \mathscr{L}$ is a separable sub-σ-algebra of $\mathscr{L}$. By theorem 1.3 there exists a set X, a σ-algebra $\mathscr{S}$ of subsets of X, and a σ-homomorphism h of $\mathscr{S}$ onto $\mathscr{L}$. Let $\{A_n : n=1, 2, \cdots\}$ be a countable family of sets of $\mathscr{S}$ such that $\mathscr{L}_1$ is the smallest sub σ-algebra of $\mathscr{L}$

containing all the $h(A_n)$. We denote by $\mathscr{S}_0$ the smallest σ-algebra of subsets of X containing all the A_n. The function

$$c : x \to (\chi_{A_1}(x), \chi_{A_2}(x), \cdots, \chi_{A_n}(x), \cdots)$$

(where χ_A denotes the function which is 1 on A, and 0 on $X-A$) is $\mathscr{S}_0$-measurable from X into the compact metric space Y which is the product of countably many copies of the 2-point space consisting of 0 and 1. Moreover, it is obvious that each A_n is of the form $c^{-1}(F)$ for some Borel set $F \subseteq Y$, and hence $\mathscr{S}_0 = \{c^{-1}(F) : F \text{ Borel in } Y\}$. Now, by a classical theorem (Kuratowski [1]), there exists a Borel isomorphism d of Y onto R^1, so that the function $c_1 : x \to d(c(x))$ is an $\mathscr{S}_0$-measurable real valued function and $\mathscr{S}_0 = \{c_1{}^{-1}(E) : E \text{ Borel in } R^1\}$. If we now define, for any Borel set E of R^1, $u(E)$ by the equation

$$u(E) = h(c_1{}^{-1}(E)),$$

then u is an observable associated with $\mathscr{L}$ whose range is $\mathscr{L}_1$. This proves (i).

We now come to the proof of (ii). Suppose $u_1, u_2, \cdots, u_n$ are observables associated with $\mathscr{L}$, having ranges $\mathscr{L}_1, \cdots, \mathscr{L}_n$, respectively. Each $\mathscr{L}_i$ is separable and hence $\mathscr{L}_0$, the smallest sub-σ-algebra of $\mathscr{L}$ containing all the $\mathscr{L}_i$, is also separable. Let X, $\mathscr{S}$, and h have the same significance as in the proof of (i). By theorem 1.4, there exists a real valued $\mathscr{S}$-measurable function f_i on X such that $u_i(E) = h(f_i{}^{-1}(E))$ for all Borel subsets E of R^1. Let $\tilde{f}$ be the map $x \to (f_1(x), \cdots, f_n(x))$ of X into R^n. Then $\tilde{f}$ is $\mathscr{S}$-measurable. The map $u : F \to h(\tilde{f}^{-1}(F))$ is then a σ-homomorphism of $\mathscr{B}(R^n)$ into $\mathscr{L}$ such that $u_i(E) = u(p_i{}^{-1}(E))$ for all $E \in \mathscr{B}(R^1)$. Since $\mathscr{B}(R^n)$ is the smallest σ-algebra of subsets of R^n containing all the sets $p_i{}^{-1}(E)$, it is clear that the range of u is $\mathscr{L}_0$. The uniqueness of u is obvious.

For any real Borel function φ on R^n, $E \to u(\varphi^{-1}(E))$ is an observable associated with $\mathscr{L}$ whose range is obviously contained in $\mathscr{L}_0$. Suppose now that v is an observable associated with $\mathscr{L}$ whose range $\mathscr{L}_v \subseteq \mathscr{L}_0$. If we use the notations of the previous paragraph, and define $\mathscr{S}^\sim$ by

$$\mathscr{S}^\sim = \{\tilde{f}^{-1}(F) : F \in \mathscr{B}(R^n)\},$$

then h maps $\mathscr{S}^\sim$ onto $\mathscr{L}_0$. Applying theorem 1.4 to $\mathscr{S}^\sim$ and v, we infer the existence of a real valued $\mathscr{S}^\sim$-measurable Borel function c on X such that $h(c^{-1}(E)) = v(E)$ for all Borel sets E of R^1. By lemma 1.5, since c is $\mathscr{S}^\sim$-measurable, there exists a real valued Borel function φ on R^n such that $c(x) = \varphi(\tilde{f}(x))$ for all $x \in X$. If now E is any Borel set on the line, we have:

$$\begin{aligned} u(\varphi^{-1}(E)) &= h(\tilde{f}^{-1}(\varphi^{-1}(E))) \\ &= h(c^{-1}(E)) \\ &= v(E). \end{aligned}$$

This proves (ii) and completes the proof of the theorem.

Remark. The uniqueness of u, guaranteed by (ii) of theorem 1.6 shows that u is independent of the constructs X, $\mathscr{S}$, and h, used in its construction. Consequently, for any real Borel function φ on R^n, the map $E \to u(\varphi^{-1}(E))$ is uniquely determined by $u_1, u_2, \cdots, u_n$ and φ. It is natural to denote this observable associated with $\mathscr{L}$ by $\varphi(u_1, u_2, \cdots, u_n)$.

NOTES ON CHAPTER I

1. For a beautiful modern introduction to classical mechanics, see the book of V. I. Arnold, *Mathematical Methods of Classical Mechanics*, Springer-Verlag, New York, 1978.

CHAPTER II

PROJECTIVE GEOMETRIES

1. COMPLEMENTED MODULAR LATTICES

The point of departure of our discussion of quantum phenomenology is the observation that the partially ordered set of experimentally verifiable statements associated with an atomic system cannot be expected to possess the distributivity properties characteristic of the Boolean algebras associated with classical systems. The simplest and most interesting of the mathematical structures that model such systems are the projective geometries, namely, the lattices of subspaces of vector spaces. More precisely, let $\mathbf{D}$ be a division ring, and let V be a vector space of finite dimension $n \geqslant 2$ over $\mathbf{D}$; we shall always suppose our vector spaces to be *left* vector spaces (cf. Jacobson [1]). Then $\mathscr{L} = \mathscr{L}(V, \mathbf{D})$, the lattice of linear subspaces of V partially ordered by inclusion, is called the *projective geometry of* V. This chapter will begin with a brief review of the basic properties of projective geometries. Although these are too simple to serve as models for realistic quantum systems, they already possess many of the fundamental features of the more complex systems.

The first example of an abstract projective geometry that anyone comes across is the projective plane, described axiomatically in terms of its points, lines, and their incidence properties. To define higher dimensional projective spaces one replaces incidence by partial order and uses a dimension function to characterize the hierarchies of points, lines, planes, and so on; the properties of incidence that are characteristic of a projective geometry are then summarized by certain natural properties of the dimension function.

We recall that a lattice $\mathscr{L}$ is *complemented* if for any $a \in \mathscr{L}$ there is a $b \in \mathscr{L}$ such that $a \wedge b = 0$ and $a \vee b = 1$. A complemented lattice $\mathscr{L}$ is *modular* if

$$(1) \qquad a, b, c \in \mathscr{L}, \qquad c < a \Rightarrow a \wedge (b \vee c) = (a \wedge b) \vee c.$$

Boolean algebras are modular; for any division ring $\mathbf{D}$ and any finite dimensional vector space V over $\mathbf{D}$, $\mathscr{L}(V, \mathbf{D})$ is a modular lattice which is not a Boolean algebra if $\dim(V) \geqslant 2$. We next introduce the notion of a lattice of finite rank. A *chain* in a lattice is a sequence $(a_i)_{1 \leqslant i \leqslant n}$ of elements a_i of the lattice such that $a_1 < a_2 < \ldots < a_n$, with $a_1 \neq 0$ and $a_i \neq a_j$ for $i < j$; n is called the *length* of the chain, and the *rank* of the lattice is the supremum of the lengths of all possible chains. An element of a lattice is called a *point* if it is nonzero and minimal. If $x_1, x_2, \ldots, x_n$ are points, they are called *independent* if $x_i \nleq \bigvee_{j \neq i} x_j$ for $i = 1, 2, \ldots, n$. The lattice $\mathscr{L}(V,\mathbf{D})$ has rank $\dim(V)$, its points are the one-dimensional subspaces, and independence is just the usual linear independence. For any lattice $\mathscr{L}$ and any $a \neq 0$ in $\mathscr{L}$, let us write $\mathscr{L}[0,a]$ for the set of all elements b of $\mathscr{L}$ with $0 < b < a$; under the partial ordering inherited from $\mathscr{L}$, $\mathscr{L}[0,a]$ is also a lattice. If $\mathscr{L}$ is of finite rank, so is $\mathscr{L}[0,a]$; and we refer to its rank as the *dimension*, $\dim(a)$, of a. We use the convention that $\dim(0) = 0$.

If $\mathscr{L}$ is a complemented modular lattice of finite rank n, it is not difficult to show that the function $a \to \dim(a)$ has the following properties:

(i) $\dim(0) = 0$, $\dim(a)$ is an integer $\geqslant 0$ for all a;
(ii) if $a < b$, $\dim(a) \leqslant \dim(b)$, with equality only for $a = b$;
(2) (iii) $\dim(a \vee b) + \dim(a \wedge b) = \dim(a) + \dim(b)$ for all a, b;
(iv) if $s = \dim(a) > 0$, then s is the maximum number of independent points contained in a; and a family of independent points contained in a has sum equal to a if and only if it has s elements.

In particular $\dim(1) = n$ and $\dim(x) = 1$ for any point x. The converse is also true; if $\mathscr{L}$ is a complemented lattice of finite rank admitting a function $d: \mathscr{L} \to \mathbf{R}$ satisfying (i)–(iii) of (2) and such that $d(x) = 1$ for all points x, then $\mathscr{L}$ is modular and $d(a) = \dim(a)$ for all a.

The lattice $\mathscr{L}(V,\mathbf{D})$ is modular, dim being the usual linear space dimension. At the other extreme is the Boolean algebra of all subsets of a finite set, dim being the cardinality. To single out the geometries we need the notion of irreducibility. An element of a complemented modular lattice $\mathscr{L}$ is called *central* if it has a unique complement in $\mathscr{L}$, which is written as a'; the reason for this terminology is the easily proved fact that $a \in \mathscr{L}$ is central if and only if for all $x \in \mathscr{L}$

$$x = (x \wedge a) \vee (x \wedge a'). \tag{3}$$

The set of all central elements is called the *center* of $\mathscr{L}$; it is a Boolean algebra, which coincides with $\mathscr{L}$ only when $\mathscr{L}$ itself is a Boolean algebra. $\mathscr{L}$ is called *irreducible* if 0 and 1 are its only central elements. A *geometry* is an irreducible complemented modular lattice of finite rank. The lattices

$$\mathscr{L}(V,\mathbf{D}) \qquad (\dim(V) \geqslant 2)$$

are geometries.

Every complemented modular lattice of finite rank may be viewed in a canonical manner as a direct sum of geometries over its centre. More precisely, let $\mathscr{L}$ be a complemented modular lattice of finite rank and let $c_1, \ldots, c_k$ be the minimal (nonzero) elements of the center of $\mathscr{L}$. Let

$$\mathscr{L}_j = \mathscr{L}[0,c_j] \qquad \text{and let} \qquad \bar{\mathscr{L}} = \mathscr{L}_1 \times \ldots \times \mathscr{L}_k$$

where the partial order in $\bar{\mathscr{L}}$ is defined by saying that

$$(x_1, \ldots, x_k) < (y_1, \ldots, y_k)$$

if and only if $x_i < y_i$ for $1 \leqslant i \leqslant k$. Then each $\mathscr{L}_j$ is a geometry, and the map that takes $x \in \mathscr{L}$ to

$$(x \wedge c_1, x \wedge c_2, \ldots, x \wedge c_k) \in \bar{\mathscr{L}}$$

is an isomorphism of $\mathscr{L}$ with $\bar{\mathscr{L}}$, the inverse map taking $(x_1, \ldots, x_k)$ to $x_1 \vee \ldots \vee x_k$. For arbitrary geometries $\mathscr{L}_j$, $\bar{\mathscr{L}}$ defined as above is a complemented modular lattice whose center has $(0, \ldots 0, 1, 0, \ldots, 0)$ (1 in the jth place) as its minimal elements $(1 \leqslant j \leqslant k)$.

The usual geometrical terminology with which we are familiar may be introduced as follows. Let $\mathscr{L}$ be a complemented modular lattice of finite rank. If $a \in \mathscr{L}$ and $\dim(a) = 1$, a is a *point*; if $\dim(a) = 2$, a is a *line*; if $\dim(a) = 3$, a is a *plane*. If $a = b \vee c$ where b and c are distinct points, a is called the line *joining* these points, and b, c are said to be *on* a; if d is a point $< a$, b, c, d are said to be *collinear*. It can be proved that $\mathscr{L}$ is a geometry if and only if every line has at least three points on it.

The fundamental theorem of classical geometry is that every geometry of rank $\geqslant 4$ is isomorphic to some $\mathscr{L}(V, \mathbf{D})$ wher the division ring $\mathbf{D}$ is determined uniquely up to isomorphism; and that if the rank is 3, this is true if (and only if) the plane in question is *Desarguesian*. We shall discuss this question a little later. At this time we shall take a closer look at the geometries $\mathscr{L}(V, \mathbf{D})$.

2. ISOMORPHISMS OF PROJECTIVE GEOMETRIES. SEMILINEAR TRANSFORMATIONS.

The first problem that arises naturally is to obtain a description of all possible isomorphisms between $\mathscr{L}(V, \mathbf{D})$ and $\mathscr{L}(V', \mathbf{D}')$. We suppose the vector spaces to be always finite dimensional and that the scalars act from the left. To formulate the answer we need the concept of semilinear transformations. Let $\mathbf{D}$, $\mathbf{D}'$ be two division rings which are isomorphic and let

$$\sigma : c \to c^\sigma \qquad (c \in \mathbf{D}) \tag{4}$$

be an isomorphism of $\mathbf{D}$ onto $\mathbf{D}'$. Let V (resp. V') be a vector space over $\mathbf{D}$ (resp. $\mathbf{D}'$). A map L $(v \to Lv)$ of V into V' is said to be *semilinear relative to* σ or *σ-linear* if

$$\begin{aligned} L(v_1 + v_2) &= Lv_1 + Lv_2 && (v_1, v_2 \in V), \\ L(cv) &= c^\sigma Lv && (c \in \mathbf{D}, v \in V). \end{aligned} \tag{5}$$

If L is in addition a bijection of V onto V', we say that L is a *σ-linear isomorphism*.

If L is a σ-linear isomorphism of V with V', then for any linear subspace $M \subset V$, its image $L[M]$ under L is a linear subspace of V', and $M \to L[M]$ is an isomorphism of $\mathscr{L}(V,\mathbf{D})$ with $\mathscr{L}(V',\mathbf{D}')$, denoted by ξ_L. We can now formulate the main result.

Theorem 2.1. *There exists an isomorphism between* $\mathscr{L}(V,\mathbf{D})$ *and* $\mathscr{L}(V',\mathbf{D}')$ *if and only if* $\mathbf{D}$ *and* $\mathbf{D}'$ *are isomorphic and* $\dim(V)=\dim(V')$. *If* ξ *is an isomorphism of* $\mathscr{L}(V,\mathbf{D})$ *with* $\mathscr{L}(V',\mathbf{D}')$, *there is an isomorphism* $\sigma(\mathbf{D} \rightrightarrows \mathbf{D}')$ *and a* σ*-linear isomorphism* $L(V \rightrightarrows V')$ *such that* $\xi = \xi_L$. *If* $\tau(\mathbf{D} \rightrightarrows \mathbf{D}')$ *is another isomorphism and* $L'(V \rightrightarrows V')$ *a* τ*-linear isomorphism, then* $\xi_L = \xi_{L'}$ *if and only if for some* $d \in \mathbf{D}$, $d \neq 0$, *we have*

$$(6) \qquad c^\tau = dc^\sigma d^{-1} \quad (c \in \mathbf{D}), \qquad L'v = dLv \quad (v \in V).$$

For the proof we refer to Baer [1]. If $\mathbf{D}=\mathbf{D}'$ it makes sense to call an isomorphism $\xi(\mathscr{L}(V,\mathbf{D}) \rightrightarrows \mathscr{L}(V,\mathbf{D}'))$ *linear* if there is a linear isomorphism $L(V \rightrightarrows V')$ such that $\xi = \xi_L$. Theorem 2.1 leads to

Theorem 2.2. *In order that every isomorphism of* $\mathscr{L}(V,\mathbf{D})$ *with* $\mathscr{L}(V',\mathbf{D})$ *be linear it is necessary and sufficient that every automorphism of* $\mathbf{D}$ *be inner.*

The division rings of greatest importance in physics are $\mathbf{R}$, the field of real numbers, $\mathbf{C}$, the field of complex numbers, and $\mathbf{H}$, the division ring of "Hamiltonian" quaternions.

(a) $\mathbf{R}$. The identity map of $\mathbf{R}$ is its only automorphism, i.e., $\mathbf{R}$ is *rigid*. So every isomorphism between real projective geometries is linear and the linear map is determined up to a multiplicative constant.

(b) $\mathbf{C}$. One knows from the theory of algebraic fields that there are infinitely many automorphisms (cf. Bourbaki [1], pp. 114–115). However, the identity and complex conjugation ($c \to c^*$) are the only analytically well-behaved ones (e.g., measurable, bounded, etc.). Since $\mathbf{C}$ is commutative, any isomorphism between two complex geometries determines uniquely the automorphism of $\mathbf{C}$ associated with it.

(c) $\mathbf{H}$, *the quaternions.* Let

$$\mathbf{1} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \qquad \mathbf{j}_1 = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}, \qquad \mathbf{j}_2 = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \qquad \mathbf{j}_3 = \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix},$$

where i is the usual square root of -1 in the complex number field. We define $\mathbf{H}$ to be the $\mathbf{R}$-linear span of $\mathbf{1}$ and $\mathbf{j}_a$ ($1 \leqslant a \leqslant 3$); since $\mathbf{j}_a{}^2 = -\mathbf{1}$ ($1 \leqslant a \leqslant 3$) and $\mathbf{j}_a\mathbf{j}_b = -\mathbf{j}_b\mathbf{j}_a = \mathbf{j}_c$ whenever (abc) is an even permutation of (123), $\mathbf{H}$ is an (associative) algebra. It is easy to verify that it is the algebra generated by $\mathbf{j}_a$ with the above relations. Let

$$|\mathbf{q}| = +\det(\mathbf{q})^{\frac{1}{2}} = +(q_0{}^2 + q_1{}^2 + q_2{}^2 + q_3{}^2)^{\frac{1}{2}}$$

where $\mathbf{q} = q_0\mathbf{1} + q_1\mathbf{j}_1 + q_2\mathbf{j}_2 + q_3\mathbf{j}_3$ ($q_0, q_a \in \mathbf{R}$). Then $|\cdot|$ is a norm on $\mathbf{H}$ which is multiplicative, i.e., $|\mathbf{q}\mathbf{q}'| = |\mathbf{q}||\mathbf{q}'|$, and $\mathbf{q}$ is invertible in $\mathbf{H}$ if and only if

$\mathbf{q} \neq \mathbf{0}$. Thus **H** is a division ring. For $\mathbf{q}\in\mathbf{H}$, say $\mathbf{q}=q_0\mathbf{1}+\sum_{1\leqslant a\leqslant 3} q_a\mathbf{j}_a$, its conjugate $\mathbf{q}^*$ is the quaternion $q_0\mathbf{1}-\sum_{1\leqslant a\leqslant 3} q_a\mathbf{j}_a$; the map $\mathbf{q}\rightarrow\mathbf{q}^*$ is an involutive anti-automorphism of **H**. A quaternion of norm 1 is called a *unit quaternion*. A quaternion is a unit quaternion if and only if its matrix is

$$\begin{pmatrix} q_0+iq_1 & -q_2-iq_3 \\ q_2-iq_3 & q_0-iq_1 \end{pmatrix}, \qquad q_0^2+q_1^2+q_2^2+q_3^2 = 1,$$

so that the unit quaternions constitute the group $SU(2,\mathbf{C})$ of 2×2 unitary matrices of determinant 1. Two quaternions are said to be in the *same class* if they can be transformed into each other by an inner automorphism; for $\mathbf{q}$, $\mathbf{q}'$ to be in the same class it is necessary and sufficient that

$$\mathrm{tr}(\mathbf{q}) = \mathrm{tr}(\mathbf{q}') \quad (\mathrm{tr} = \text{trace}) \quad \text{and} \quad |\mathbf{q}| = |\mathbf{q}'|.$$

All the automorphisms of H are inner so that isomorphisms of quaternionic projective geometries are linear. The map $x\rightarrow x\mathbf{1}$ $(x\in\mathbf{R})$ imbeds **R** as the center of **H**.

H may be viewed as a vector space over **C**, although in a noncanonical manner. We identify **C** as a subfield of **H** by the correspondence

$$a_0+ia_1 \leftrightarrow a_0+a_1\mathbf{j}_1;$$

H is then a vector space over **C**, scalar multiplication being from the left. Then $\dim_{\mathbf{C}}(\mathbf{H})=2$, and $\{\mathbf{1},\mathbf{j}_2\}$ may be taken as a basis. For any $\mathbf{q}\in\mathbf{H}$, $\mathbf{R}_{\mathbf{q}}:\mathbf{q}'\rightarrow\mathbf{q}'\mathbf{q}^*$ is then an endomorphism of **H** over **C**, $\mathbf{q}\rightarrow\mathbf{R}_{\mathbf{q}}$ is a faithful representation, and the matrix of $\mathbf{j}_a$ $(1\leqslant a\leqslant 3)$ is $\epsilon_a\mathbf{j}_a$, where $\epsilon_a=-1$ for $a=1,3$ and $+1$ for $a=2$.

3. DUALITIES AND POLARITIES

We shall now examine the anti-automorphisms and orthocomplementations of the projective geometries $\mathscr{L}(V,\mathbf{D})$. It will turn out that orthocomplementations essentially arise only from Hilbert space structures, at least when **D** is one of **R**, **C**, and **H**.

We begin with the concept of the dual of a division ring. Let **D** be a division ring. We define $\mathbf{D}^0$ to be the ring with the same elements as **D**, with the same rule for addition, but reversing the order of multiplication. $\mathbf{D}^0$ is a division ring, said to be *dual* to **D**. Isomorphisms of $\mathbf{D}_1$ with $\mathbf{D}_2{}^0$ may be viewed as anti-isomorphisms of $\mathbf{D}_1$ with $\mathbf{D}_2$. If V is a vector space over **D**, its dual, defined as the space of **D**-linear maps of V into **D**, becomes a vector space over $\mathbf{D}^0$ if we define scalar multiplication by writing

$$(a\cdot f)(v) = f(v)a \quad (a\in\mathbf{D}^0, v\in V, f\in V^*);$$

addition in V^* is defined in the usual way. V^* is then a (left) vector space over $\mathbf{D}^0$ and has the same dimension as V. Clearly $(\mathbf{D}^0)^0=\mathbf{D}$ and so V can be canonically identified with V^{**}. If $(v_1, \ldots, v_n)$ is a basis of V, the corresponding dual basis $(v_1^*, \ldots, v_n^*)$ of V^* is defined by $v_j^*(v_k) = \delta_{jk}$. If $M \subset V$ is a linear subspace, its *annihilator* is the linear subspace of V^* defined by

$$M^0 = \{f : f \in V^*, f(v) = 0 \quad \text{for all } v \in M\}. \tag{7}$$

It is clear that

$$\dim(M) + \dim(M^0) = \dim(V) \tag{8}$$

and so $M = M^{00}$. The map $M \to M^0$ is obviously an *inclusion-reversing* bijection of $\mathscr{L}(V,\mathbf{D})$ with $\mathscr{L}(V^*,\mathbf{D}^0)$. We call it a *duality*, and use the same term to refer to any inclusion-reversing bijection of one geometry onto another.

To determine the most general duality we need the concept of a nonsingular semibilinear form. Let θ be an anti-automorphism of $\mathbf{D}$, V a vector space over $\mathbf{D}$. A *θ-bilinear form* is a map $\langle\,.\,,\,.\,\rangle : x, y \to \langle x, y\rangle$ of $V \times V$ into $\mathbf{D}$ with the following properties:

$$\text{(9)}\quad \text{(i)}\quad \begin{aligned} \langle x_1 + x_2, y\rangle &= \langle x_1, y\rangle + \langle x_2, y\rangle \\ \langle x, y_1 + y_2\rangle &= \langle x, y_1\rangle + \langle x, y_2\rangle \end{aligned} \qquad (x, y, x_1, y_1, x_2, y_2 \in V);$$

$$\text{(ii)}\quad \langle cx, dy\rangle = c\langle x, y\rangle d^\theta \qquad (c, d \in \mathbf{D},\ x, y \in V).$$

We call $\langle\,.\,,\,.\,\rangle$ *nonsingular* if

$$\langle x, y\rangle = 0 \qquad \forall\, x \in V \Rightarrow y = 0 \tag{10r}$$

and

$$\langle x, y\rangle = 0 \qquad \forall\, y \in V \Rightarrow x = 0. \tag{10l}$$

For any anti-automorphism θ of $\mathbf{D}$, the map

$$(x_1, \ldots, x_n), (y_1, \ldots, y_n) \to x_1 y_1^\theta + \ldots + x_n y_n^\theta$$

is a nonsingular θ-bilinear form on $\mathbf{D}^n \times \mathbf{D}^n$.

Theorem 2.3. *Let V be a vector space of dimension n over $\mathbf{D}$; θ, an anti-automorphism of $\mathbf{D}$; and $\langle\,.\,,\,.\,\rangle$, a θ-bilinear form on $V \times V$. Then $\langle\,.\,,\,.\,\rangle$ is nonsingular if and only if it satisfies either of* (10l) *or* (10r). *For any linear subspace M of V let*

$$M' = \{u : u \in V, \langle x, u\rangle = 0 \text{ for all } x \in M\}. \tag{11}$$

Then, if $\langle\,.\,,\,.\,\rangle$ is nonsingular, $M \to M'$ is a duality of $\mathscr{L}(V,\mathbf{D})$. If $n \geqslant 3$, every duality of $\mathscr{L}(V,\mathbf{D})$ arises in this manner, for suitable θ and $\langle\,.\,,\,.\,\rangle$. The pair θ', $\langle\,.\,,\,.\,\rangle'$ induces the same duality as θ, $\langle\,.\,,\,.\,\rangle$ if and only if there is a nonzero $d \in \mathbf{D}$ such that for all $x, y \in V$, $c \in \mathbf{D}$,

$$\langle x, y\rangle' = \langle x, y\rangle d, \qquad c^{\theta'} = d^{-1} c^\theta d. \tag{12}$$

In particular, if $\mathbf{D}$ is commutative, $\theta' = \theta$.

Proof. For $y \in V$ let $t_y \in V^*$ be defined by $t_y(u) = \langle u, y\rangle$ $(u \in V)$. Then t $(y \to t_y)$ is a θ-linear map of V into V^*. Clearly (10r) is equivalent to the statement that t is injective while (101) is equivalent to the statement that (0) is the annihilator in V of the range of t; i.e., that t is surjective. The first statement is now clear. Suppose now that $\langle .\,,.\rangle$ is nonsingular. Then t is a θ-linear isomorphism and so, as $M'=t^{-1}(M^0)$ we see that $M\to M'$ is a duality of $\mathscr{L}(V,\mathbf{D})$. Conversely, if ξ is a duality of $\mathscr{L}(V,\mathbf{D})$, $\eta: M\to(\xi^{-1}(M))^0$ is an isomorphism of $\mathscr{L}(V,\mathbf{D})$ onto $\mathscr{L}(V^*,\mathbf{D}^0)$. This shows already that

$$\dim(M)+\dim(\xi(M)) = n. \tag{13}$$

Moreover, theorem 2.1 gives the existence of an anti-automorphism θ of $\mathbf{D}$ and a θ-linear isomorphism t of V with V^* such that $M^0=t[\xi(M)]$ for all M in $\mathscr{L}(V,\mathbf{D})$. If we set $\langle x,y\rangle=(ty)(x)$ $(x,y\in V)$, it is immediate that $\langle .\,,.\rangle$ is a nonsingular θ-bilinear form and $\xi(M)=M'$ for all M. If θ' and $\langle .\,,.\rangle'$ is another pair and $t_y'(x)=\langle x,y\rangle'$ $(x,y\in V)$, θ', $\langle .\,,.\rangle'$ also give rise to ξ if and only if t and t' induce the same isomorphism of $\mathscr{L}(V,\mathbf{D})$ with $\mathscr{L}(V^*,\mathbf{D}^0)$. Theorem 2.1 now leads to the required result. This completes the proof of theorem 2.3.

A duality ξ of $\mathscr{L}(V,\mathbf{D})$ which is involutive is known as a *polarity*; this means $M=\xi(\xi(M))$ for all M. A polarity is called *isotropic* if $M\subseteq\xi(M)$ for all *one-dimensional* M.

Theorem 2.4. *Let V be a vector space of dimension $n\geqslant 3$ over $\mathbf{D}$. Then $\mathscr{L}(V,\mathbf{D})$ admits an isotropic polarity if and only if $\mathbf{D}$ is commutative and* $\dim(V)$ *is even. In this case, if ξ is an isotropic polarity and $2N=\dim(V)$, we can find a nonsingular skew-symmetric bilinear form $\langle .\,,.\rangle$ such that $\xi(M)=M'$ for all M. Moreover, we can find a basis $\{x_1,y_1,\ldots,x_N,y_N\}$ for V such that $\langle x_i,x_j\rangle=\langle y_i,y_j\rangle=0$, $\langle x_i,y_j\rangle=-\langle y_j,x_i\rangle=\delta_{ij}$ $(1\leqslant i,j\leqslant N)$. In particular, if ξ and ξ' are two isotropic polarities of $\mathscr{L}(V,\mathbf{D})$, there is a linear automorphism a of $\mathscr{L}(V,\mathbf{D})$ such that $\xi'=a\xi a^{-1}$.*

Proof. Let ξ be an isotropic polarity of $\mathscr{L}(V,\mathbf{D})$. By theorem 2.3 there exists an anti-automorphism θ of $\mathbf{D}$ and a nonsingular θ-bilinear form $\langle .\,,.\rangle$ giving rise to ξ. Since ξ is isotropic $\langle x,x\rangle=0$ for all $x\in V$. Replacing x by $x+y$ we see that $\langle .\,,.\rangle$ is skew symmetric. If now $c\in\mathbf{D}$,

$$\langle x,y\rangle c^\theta = -c\langle y,x\rangle = c\langle x,y\rangle;$$

choosing x, y such that $\langle x,y\rangle=1$ we see that θ is the identity. This means that $\mathbf{D}$ is commutative and $\langle .\,,.\rangle$ is bilinear. The remaining statements are standard. Theorem 2.4 is proved.

We now consider nonisotropic polarities. A θ-bilinear form $\langle .\,,.\rangle$ is called *symmetric* if $\langle x,y\rangle^\theta=\langle y,x\rangle$ for all $x,y\in V$. Notice that this property is relative to θ. We observe that if $\langle .\,,.\rangle$ is not identically zero, θ is necessarily involutive; then the set of values of $\langle .\,,.\rangle$ is all of $\mathbf{D}$, and

$$\langle x,y\rangle^{\theta^2} = \langle y,x\rangle^\theta = \langle x,y\rangle.$$

Lemma 2.5. *Let* $\dim(V)\geqslant 2$, $\langle\,.\,,\,.\,\rangle$ *a nonsingular* θ*-bilinear form such that* $\langle x,x\rangle=\langle x,x\rangle^{\theta}\neq 0$ *for some* $x\in V$. *Suppose* $\langle\,.\,,\,.\,\rangle$ *induces a polarity. Then this polarity is nonisotropic,* θ *is involutive, and* $\langle\,.\,,\,.\,\rangle$ *is symmetric.*

Proof. If $u,v\in V,\langle u,v\rangle=0\Leftrightarrow \mathbf{D}\cdot v\subseteqq(\mathbf{D}\cdot u)'\Leftrightarrow(\mathbf{D}\cdot u)\subseteqq(\mathbf{D}\cdot v)'\Leftrightarrow\langle v,u\rangle=0$. Let $c=\langle x,x\rangle$. Then $x\notin(\mathbf{D}\cdot x)'$ and $\dim((\mathbf{D}\cdot x)')=n-1$, so that

$$V=(\mathbf{D}\cdot x)\oplus(\mathbf{D}\cdot x)'.$$

If $u,v\in(\mathbf{D}\cdot x)'$, $\langle\langle u,v\rangle x-u,x+cv\rangle=0$, so that $\langle x+cv,\langle u,v\rangle x-u\rangle=0$, proving that $\langle u,v\rangle^{\theta}=\langle v,u\rangle$. Now $\langle\,.\,,\,.\,\rangle$ is not identically zero on

$$(\mathbf{D}.x)'\times(\mathbf{D}.x)',$$

and so this already proves that θ is involutive. The symmetry of $\langle\,.\,,\,.\,\rangle$ on $V\times V$ now follows from direct computation. Lemma 2.5 is proved.

Theorem 2.6. *Let* θ *be an involutive anti-automorphism of* $\mathbf{D}$, V *a vector space of dimension* n *over* $\mathbf{D}$, *and* $\langle\,.\,,\,.\,\rangle$ *a nonsingular symmetric* θ*-bilinear form on* $V\times V$. *Then the duality corresponding to* $\langle\,.\,,\,.\,\rangle$ *is a nonisotropic polarity unless* $\mathbf{D}$ *is commutative and of characteristic* 2. *Conversely, let* $n\geqslant 3$ *and let* ξ *be a nonisotropic polarity of* $\mathscr{L}(V,\mathbf{D})$. *Then there exists an involutive anti-automorphism* θ *of* $\mathbf{D}$ *and a nonsingular symmetric* θ*-bilinear form* $\langle\,.\,,\,.\,\rangle$ *on* $V\times V$ *such that* ξ *is induced by* $\langle\,.\,,\,.\,\rangle$. *If* θ' *and* $\langle\,.\,,\,.\,\rangle'$ *is another pair,* ξ *is also induced by them if and only if there is a* $d\neq 0$ *in* $\mathbf{D}$ *such that for all* $x,y\in V$, $c\in\mathbf{D}$

$$\langle x,y\rangle'=\langle x,y\rangle d,\qquad c^{\theta'}=d^{-1}c^{\theta}d.$$

In this case $d^{\theta}=d$.

Proof. Given θ and $\langle\,.\,,\,.\,\rangle$, the corresponding duality is a polarity. If this were isotropic we can argue as in theorem 2.4 to conclude that $\mathbf{D}$ is commutative, $\theta=$ identity, and $\langle\,.\,,\,.\,\rangle$ is skew symmetric. As $\langle\,.\,,\,.\,\rangle$ is symmetric it follows that $\mathbf{D}$ has characteristic 2. Conversely, let $n\geqslant 3$ and let ξ be a nonisotropic polarity of $\mathscr{L}(V,\mathbf{D})$. Then ξ corresponds to a pair θ_0, $\langle\,.\,,\,.\,\rangle$. Choose $x\in V$ such that $\langle x,x\rangle_0=c_0\neq 0$. Writing

$$\langle u,v\rangle=\langle u,v\rangle_0 c_0^{-1},\qquad c^{\theta}=c_0c^{\theta}c_0^{-1},$$

we see that ξ is also induced by θ, $\langle\,.\,,\,.\,\rangle$, and that $\langle x,x\rangle=1$. Lemma 2.5 shows that θ is involutive, and $\langle\,.\,,\,.\,\rangle$ is symmetric; clearly, θ, $\langle\,.\,,\,.\,\rangle$ induce ξ. Suppose θ' is an anti-automorphism of $\mathbf{D}$ and $\langle\,.\,,\,.\,\rangle'$ a symmetric θ'-bilinear form. Then, by theorem 2.3, θ', $\langle\,.\,,\,.\,\rangle'$ induce ξ if and only if for some $d\neq 0$ in $\mathbf{D}$, $\langle x,y\rangle'=\langle x,y\rangle d$, $(x,y\in V)$ and $c^{\theta'}=d^{-1}c^{\theta}d$ for all $c\in\mathbf{D}$. If $\langle x,y\rangle=1$, then $\langle y,x\rangle=1$ and so $(\langle x,y\rangle')^{\theta'}=\langle y,x\rangle'=d$ while

$$(\langle x,y\rangle')^{\theta'}=d^{\theta'}=d^{-1}d^{\theta}d,$$

also. So we have $d^{\theta}=d$.

Remark. It is useful to note that the form $\langle .\,,. \rangle$ inducing ξ has the property that $\langle w,w \rangle = 1$ for some $w \in V$.

4. ORTHOCOMPLEMENTATIONS AND HILBERT SPACE STRUCTURES

An *orthocomplementation* of $\mathscr{L}(V,\mathbf{D})$ is a polarity ξ such that $M \cap \xi(M) = 0$ for all M, i.e., an inclusion reversing bijection ξ of $\mathscr{L}(V,\mathbf{D})$ with itself such that $\xi(\xi(M)) = M$ and $M \cap \xi(M) = 0$ for all M. It is easy to see that a polarity is an orthocomplementation if and only if $(\mathbf{D}\cdot x) \cap \xi(\mathbf{D}\cdot x) = 0$ for all $x \in V$. The preceding discussion leads to

Theorem 2.7 (*Birkhoff–von Neumann* [1]). *Let* $n = \dim(V) \geqslant 3$ *and* ξ *an orthocomplementation of* $\mathscr{L}(V,\mathbf{D})$. *Then there is an involutive anti-automorphism* θ *of* $\mathbf{D}$ *and a nonsingular symmetric* θ*-bilinear form* $\langle .\,,. \rangle$ *on* $V \times V$ *such that* θ *and* $\langle .\,,. \rangle$ *induce* ξ. *We have* $\langle x,x \rangle = 0 \Leftrightarrow x = 0$, *and* $\langle .\,,. \rangle$ *can be chosen so that* $\langle w,w \rangle = 1$ *for some* $w \in V$.

Let a symmetric θ-bilinear form $\langle .\,,. \rangle$ be called *definite if*

$$\text{(14)}\qquad \begin{aligned} &\text{(i)}\quad \langle x,x \rangle = 0 \;\Leftrightarrow\; x = 0,\\ &\text{(ii)}\quad \langle w,w \rangle = 1 \quad \text{for some} \quad w \in V. \end{aligned}$$

A definite form is necessarily nonsingular. Assume now that $\mathbf{D}$ is one of $\mathbf{R}$, $\mathbf{C}$, or $\mathbf{H}$.

(a) $\mathbf{D} = \mathbf{R}$. Then the definite forms are precisely the positive definite quadratic forms. So the definite form $\langle .\,,. \rangle$ of theorem 2.7 defines a scalar product and ξ is the orthocomplementation in the corresponding Hilbert space structure.

(b) $\mathbf{D} = \mathbf{C}$. Here we must assume that θ is continuous or measurable (analytically well behaved). Since complex bilinear forms are not definite, θ must be the complex conjugation. The definite forms are thus the Hermitian positive definite forms, and in theorem 2.7, ξ is the orthocomplementation defined by the Hilbert space structure corresponding to the scalar product $\langle .\,,. \rangle$.

(c) $\mathbf{D} = \mathbf{H}$, *the quaternions*. A *scalar product* for a vector space V over $\mathbf{H}$ is a *-bilinear form on $V \times V$ such that: (i) for any $x \to V$, $\langle x,x \rangle$ is in the center of $\mathbf{H}$ and coincides with a nonnegative real number; (ii) $\langle x,x \rangle = 0$ if and only if $x = 0$; if $\|x\| = +\langle x,x \rangle^{\frac{1}{2}}$, $\|\cdot\|$ is a norm for V. If V is complete under $\|\cdot\|$, it is called a Hilbert space; this is the case when $\dim(V) < \infty$ (cf. Finkelstein and others [1]). Let us now consider θ and $\langle .\,,. \rangle$ as in theorem 2.7. Then we can find a unit quaternion $\mathbf{q}_0$ such that $\mathbf{c}^\theta = \mathbf{q}_0\mathbf{c}^*\mathbf{q}_0^{-1}$ for all $\mathbf{c} \in \mathbf{H}$. Since θ is involutive, $\mathbf{q}_0^2$ lies in the center of $\mathbf{H}$ and so $\mathbf{q}_0^2 = \pm 1$.

We shall now show that if $\mathbf{q}_0^2 = -1$ there are vectors $x \neq 0$ with $\langle x,x \rangle = 0$. In fact, we can then write $\mathbf{q}_0 = \mathbf{q}_0'\mathbf{j}_1\mathbf{q}_0'^{-1}$ where $\mathbf{q}_0'$ is a unit quaternion. Let

$u \in V$ be such that $\langle u,u \rangle$ is a unit quaternion, say $\mathbf{a}$, and let $\overline{u} \in \xi\,(\mathbf{H} \cdot u)$ be such that $\langle \overline{u},\overline{u} \rangle$ is also a unit quaternion, say $\mathbf{b}$; of course $V = \mathbf{H} \cdot u \oplus \xi(\mathbf{H} \cdot u)$. From $\langle u,u \rangle^{\theta} = \langle u,u \rangle$ we get, after a small calculation,

$$\mathbf{q}_0{}'^{-1}\mathbf{a}\mathbf{q}_0{}' = \begin{pmatrix} c & r \\ -r^* & c \end{pmatrix}$$

where $c \in \mathbf{R}$, $|c^2| + |r|^2 = 1$. Hence $\mathrm{tr}(\mathbf{a}\mathbf{q}_0) = 0$. Similarly, $\mathrm{tr}(\mathbf{b}\mathbf{q}_0) = 0$. So $\mathbf{a}\mathbf{q}_0$ and $-\mathbf{b}\mathbf{q}_0$ are in the same class, proving that $\mathbf{q}_1\mathbf{a}\mathbf{q}_0\mathbf{q}_1{}^*\mathbf{q}_0{}^{-1} + \mathbf{b} = 0$ for some unit quaternion $\mathbf{q}_1$. But then $\langle x,x \rangle = 0$ for $x = \mathbf{q}_1 u + v$. So, $\mathbf{q}_0{}^2 = +\mathbf{1}$ is the only possibility. Since $\det(\mathbf{q}_0) = 1$, we must have $\mathbf{q}_0 = \pm 1$. So θ must be conjugation and $\langle\, .\, ,\, .\, \rangle$ is $*$-bilinear. In particular, $\langle u,u \rangle$ is real for all $u \in V$, and >0 for $u \in V \backslash (0)$. V thus becomes a Hilbert space under $\langle\, .\, ,\, .\, \rangle$, and ξ is the associated orthocomplementation. We thus have

Theorem 2.8. *Let* $\mathbf{D} = \mathbf{R}$ *or* $\mathbf{H}$ *and let* V *be a vector space of dimension* $n \geqslant 3$ *over* $\mathbf{D}$. *If* ξ *is any orthocomplementation in* $\mathscr{L}(V,\mathbf{D})$, *there is a scalar product for* $\langle\, .\, ,\, .\, \rangle$ *converting* V *into a Hilbert space such that* ξ *is the corresponding orthocomplementation. If* $\mathbf{D} = \mathbf{C}$ *and* ξ *is regular in the sense that the anti-automorphism of* $\mathbf{C}$ *defined by* ξ *is continuous, then it is the complex conjugation, and* ξ *is induced by a Hermitian scalar product* $\langle\, .\, ,\, .\, \rangle$ *that converts* V *into a Hilbert space over* $\mathbf{C}$. *Finally* $\langle\, .\, ,\, .\, \rangle$ *is determined by* ξ *up to multiplication by a positive real number.*

We shall now assume that $\mathbf{D}$ is one of $\mathbf{R}$, $\mathbf{C}$, or $\mathbf{H}$, and that V is a Hilbert space of dimension $n \geqslant 3$ over $\mathbf{D}$, with scalar product $\langle\, .\, ,\, .\, \rangle$.

Theorem 2.9. *Suppose* ξ *is an automorphism of* $\mathscr{L}(V,\mathbf{D})$ *which preserves the orthocomplementation of* V. *Then there exists a semilinear automorphism* L *of* V *inducing* ξ *such that*:

(i) *if* $\mathbf{D} = \mathbf{R}$ *or* $\mathbf{H}$, L *is linear, and for* $\mathbf{D} = \mathbf{C}$, L *is either linear or conjugate linear;*

(ii) *for all* $x, y \in V$,

(15) $$\langle Lx,Ly \rangle = \langle x,y \rangle$$

if L *is linear, and*

(16) $$\langle Lx,Ly \rangle = \langle y,x \rangle$$

if $\mathbf{D} = \mathbf{C}$ *and* L *is conjugate linear. In particular,* L *is an isometry of* V.

Proof. We begin with the following observation. Let φ be an automorphism of $\mathbf{D}$ and L_1 a φ-linear isomorphism of V onto itself such that the induced automorphism of $\mathscr{L}(V,\mathbf{D})$ preserves the orthocomplementation of V; then, for some $d \neq 0$ in $\mathbf{D}$, we have, for all $x, y \in V$,

(17) $$\langle L_1x,L_1y \rangle = \langle x,y \rangle^{\varphi} d$$

Indeed, $\langle x,y\rangle = 0$ if and only if $\langle L_1x,L_1y\rangle = 0$; so, if we define

$$\langle x,y\rangle^{\sim} = \langle L_1x,L_1y\rangle^{(\varphi^{-1})}\ (x, y \in V) \quad \text{and} \quad \psi = \varphi^{-1}\theta\varphi$$

where θ is the anti-automorphism corresponding to $\langle . , . \rangle$, then the pair ψ, $\langle . , . \rangle^{\sim}$ also induces the orthocomplementation in $\mathscr{L}(V,\mathbf{D})$. Theorem 2.3 now leads to (17). This said, let us first suppose that ξ is induced by a linear automorphism L_1 of V. Then (17) becomes $\langle L_1x,L_1y\rangle = \langle x,y\rangle d$. This shows that d is a real positive number, and we get the required result on taking $L = d^{-\frac{1}{2}}L_1$. If ξ is not linear, then $\mathbf{D} = \mathbf{C}$ and there is an automorphism φ of $\mathbf{C}$ which is not the identity and a φ-linear isomorphism L_1 of V with itself, such that L_1 induces ξ. In (17) we take $x = y = x_0$ where x_0 is such that $\langle x_0,x_0\rangle = 1$, so that d is a positive number. Then, taking $x = y$ in (17) shows that $\langle x,x\rangle^{\varphi}$ is real and $\geqslant 0$ for all $x \in V$, so that $r^{\varphi} \geqslant 0$ for all real numbers $\geqslant 0$. In other words $\mathbf{R}^{\varphi} = \mathbf{R}$, which implies $\varphi =$ identity on $\mathbf{R}$. As φ is not the identity, $i^{\varphi} = -i$ and so φ is complex conjugation. If $L = d^{-\frac{1}{2}}L_1$, we get (16). This proves theorem 2.9.

When $\mathbf{D} = \mathbf{C}$ and L is a conjugate linear isomorphism of V with itself such that (16) is satisfied, we call L *anti-unitary*. Of course, if L is linear and satisfies (15), L is *unitary*.

5. COORDINATES IN PROJECTIVE AND GENERALIZED GEOMETRIES

The study of the projective geometries $\mathscr{L}(V,\mathbf{D})$ gains significance in view of the fundamental theorem of classical geometry which asserts that these are the only geometries, at least when the rank is $\geqslant 4$. When the rank is 3, i.e., when we are dealing with a projective plane, the geometries $\mathscr{L}(V,\mathbf{D})$ are characterized by the property of being *Desarguesian*, namely, that the classical theorem of Desargues is true for them. Of course a projective plane imbedded in a geometry of higher rank is always Desarguesian; this can be proved directly and in a simple manner.

However, as was mentioned at the beginning of the chapter, the geometries $\mathscr{L}(V,\mathbf{D})$ and, more generally, lattices of finite rank, are not adequate to serve as models for the proposition calculi of complex quantum systems. We shall therefore enlarge the scope of our discussion by including a class of partially ordered sets of infinite rank, and proving for them a coordinization theorem that will include the classical result. The objects that we shall study will be called generalized geometries; and the main result concerning them is that they are isomorphic to the partially ordered sets of *finite dimensional* subspaces of possibly infinite dimensional vector spaces over division rings.

Let $\mathscr{L}$ be a set with partial ordering $<$; *we do not assume that $\mathscr{L}$ has a unit element.* $\mathscr{L}$ is called a *generalized geometry* if the following conditions are satisfied:

$$\text{(18)}\qquad \begin{aligned} &\text{(i)}\quad \text{for any finite subset } F \text{ of } \mathscr{L},\ \textstyle\bigvee_{a\in F} a,\ \bigwedge_{a\in F} a, \text{ exist in } \mathscr{L},\\ &\text{(ii)}\quad \text{if } a\in\mathscr{L} \text{ and } a\neq 0,\ \mathscr{L}[0,a] \text{ is a geometry.}\end{aligned}$$

Thus $\mathscr{L}$ is a geometry if and only if it has a unit element. If $\mathbf{D}$ is a division ring and V a not necessarily finite dimensional vector space over $\mathbf{D}$, the partially ordered set $\mathscr{L}(V,\mathbf{D})$ of finite dimensional subspaces of V is a generalized geometry, which is a geometry if and only if $\dim(V)<\infty$. If $\mathscr{L}$ is a generalized geometry, $\mathscr{L}$ has points and every element of $\mathscr{L}$ is a sum of points. The dimension functions of the $\mathscr{L}[0,a]$ are mutually compatible and define a dimension function on all of $\mathscr{L}$. We write $\dim(\mathscr{L})$ for

$$\sup_{a\in\mathscr{L}}(\dim(a)).$$

We can obviously speak of lines, planes, etc., in $\mathscr{L}$. In the rest of this section we shall sketch a proof that any generalized geometry is isomorphic to some $\mathscr{L}(V,\mathbf{D})$, provided $\dim(\mathscr{L})\geqslant 4$. Our proof is a straightforward variant of the classical one; but we have decided to include a sketch since it is not easy to locate it in the literature in the precise form we need. The reader who wants more details on the part dealing with Desarguesian planes should consult Seidenberg [1] and Heyting [1] (see the Notes at the end of this chapter).

Let $\mathscr{L}$ be a generalized geometry with $\dim(\mathscr{L})\geqslant 4$. Then all the planes in $\mathscr{L}$ are Desarguesian. The classical method of constructing the division rings associated to $\mathscr{L}$ is as follows. We take a line t in $\mathscr{L}$ and a plane a in $\mathscr{L}$ containing t, and we choose three distinct points O, E, W on t. To this data we shall associate a division ring $\mathbf{D}$ whose elements will be the points of the line t that are different from W. The operations of addition and multiplication in $\mathbf{D}$ will be defined with the help of geometrical constructions (in the plane a) that imitate the familiar constructions in the Euclidean plane.

Addition. Let $A\in\mathbf{D}$; then we define the map $a_A\colon B\to B+A$ of $\mathbf{D}$ with itself in the following manner. In the plane a we take distinct lines q, r through W, different from t, and distinct points X, Y on q, different from W, such that if $\pi(X; r,t)$ (resp. $\pi(Y; t,r)$ is the perspectivity from r to t with center X (resp. from t to r with center Y), then $\pi(X; r,t)\,\pi(Y; t,r)$ takes 0 to A; then we define (Fig. 1)

$$a_A=\pi(X;r,t)\,\pi(Y;t,r).$$

If we view q as the "line at infinity" then this construction resembles the usual Euclidean one (Fig. 1(a)).

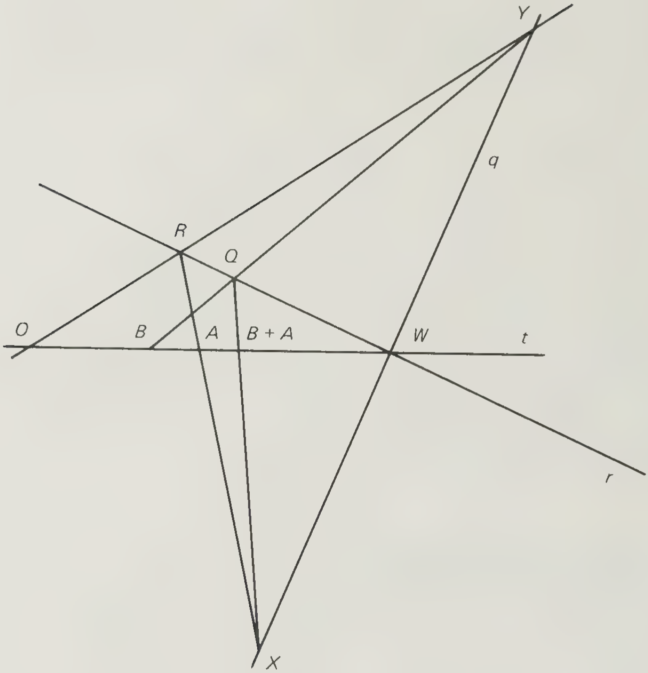

FIG. 1

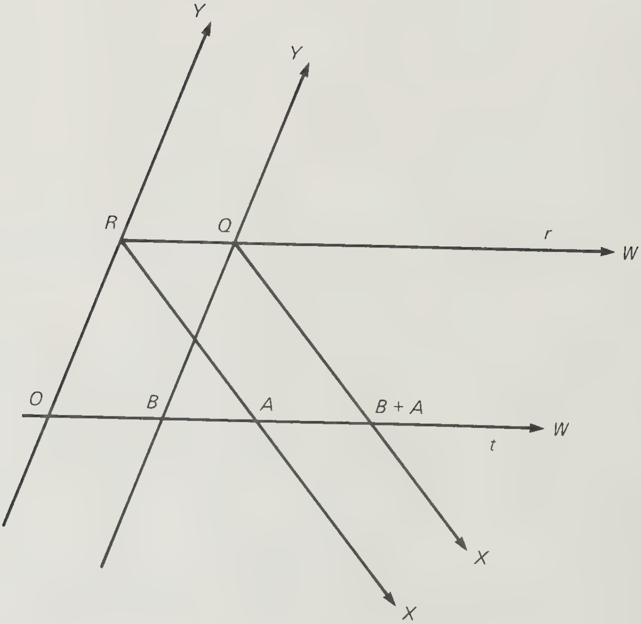

FIG. 1(a)

Multiplication. Given $A \in \mathbf{D}$ we define *right multiplication* m_A by A; if $A = O$, $m_A B = O$ by definition for all B and so we take $A \neq O$. We take in the plane a a line q through W different from t, and a line p through O, different from t. We choose distinct points X, Y on q different from W such that $\pi(X; p,t)\,\pi(Y; t,p)$ takes E to A (see Fig. 2); then

$$m_A = \pi(X; p,t)\,\pi(Y; t,p).$$

If we view q as the "line at infinity", this resembles the usual construction in the Euclidean plane (Fig. 2(a)). Of course we could have equally well taken this as the definition of $A \cdot B$; our choice is dictated by the requirement of getting isomorphisms with the geometry of subspaces of a *left* vector space.

Although these definitions involve the choices of auxiliary lines and points, the Desarguesian nature of the plane can be used to prove that addition and multiplication are well defined and convert $\mathbf{D}$ into a division ring with O and E as its null and unit elements. If t, O, E, W are fixed, this division ring is independent of the choice of the plane a that contains t; for, if b is a different plane containing t, and X is a point in the three-dimensional space $a \vee b$ which does not lie on a or b, the perspectivity from X gives an

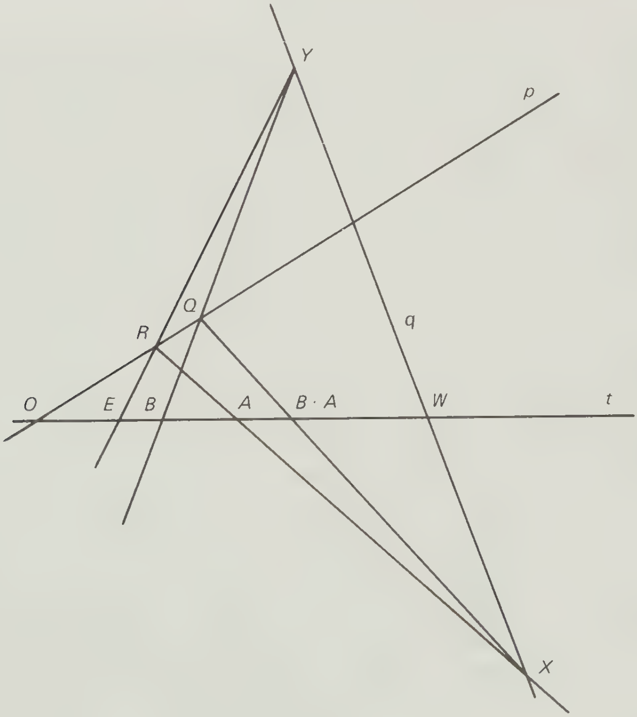

FIG. 2

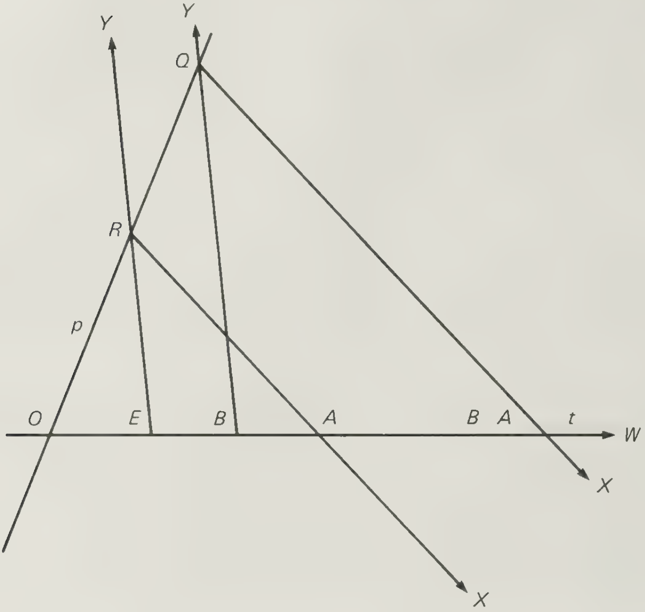

FIG. 2(a)

isomorphism of $\mathscr{L}[0,a]$ with $\mathscr{L}[0,b]$ which is the identity on $t = a \wedge b$, and it is clear that addition and multiplication on $\mathbf{D}$ are the same whether we use a or b. The division ring is also unchanged up to isomorphism if we change t O, E, W. More precisely, let t' be another line in the plane a, and O', E', W', three distinct points on t'. Then any projectivity in the plane that takes t to t' and O, E, W, respectively to O', E', W', is an isomorphism of the associated division rings; this follows from the well-known fact that any projectivity of t with t' extends to an automorphism of the plane $\mathscr{L}[0,a]$. We remark finally that $\mathbf{D}$ is commutative if and only if the geometry of $\mathscr{L}[0,a]$ is Pappian for any plane a in $\mathscr{L}$, i.e., the theorem of Pappus is valid in $\mathscr{L}[0,a]$ for any plane a.

The first step in the coordinatization of $\mathscr{L}$ is to define "affine" coordinates for all points of $\mathscr{L}$ that do not "lie at infinity". If $\mathscr{L}$ is a geometry we can choose an arbitrary hyperplane to be at infinity. For a generalized geometry we use a simple modification of this procedure. To obtain an intuitive picture of the definition of coordinates the reader is asked to picture everything in a Euclidean context, interpreting incidences at infinity in terms of parallellism.

Let us call a set of points of $\mathscr{L}$ *independent* if every finite subset of the set is independent. If $\{P_j\}$ $(j \in J)$ is an independent set of points, then we have a map u $(K \to u(K))$ from finite subsets of J into $\mathscr{L}$ defined by

$$u(K) = \bigvee_{j \in K} P_j, \qquad u(\phi) = 0. \tag{19}$$

It is easy to verify that

$$(20)\quad u(K_1 \cup K_2) = u(K_1) \vee u(K_2), \qquad u(K_1 \cap K_2) = u(K_1) \wedge u(K_2).$$

An independent set C of points of $\mathscr{L}$ is called a *basis* if it is maximal. Every independent set is contained in a basis in view of Zorn's lemma. If C is a basis of $\mathscr{L}$ and P is a point of $\mathscr{L}$, it follows from the maximality of C that there is a finite set $F \subset C$ such that $P < \bigvee_{Q \in F} Q$; hence for any $a \in \mathscr{L}$, there is a finite set $G \subset C$ with $a < \bigvee_{Q \in G} Q$. Let O be a point of $\mathscr{L}$. By a *frame* at O we mean a pair $(O, \{P_j\}_{j \in J})$ such that $\{O, \{P_j\}_{j \in J}\}$ is a basis for $\mathscr{L}$. There are frames at any point of $\mathscr{L}$.

From now on we fix a frame $(O, \{P_j\}_{j \in J})$ at O. An element $a \in \mathscr{L}$ is said to *lie at infinity* if there is a finite set $K \subset J$ such that $a < u(K)$ (cf. (19)). It follows from (20) that the elements of $\mathscr{L}$ that lie at infinity form a generalized geometry, denoted by $\mathscr{L}_\infty$. If $a \in \mathscr{L}$ and $a \notin \mathscr{L}_\infty$, there is a unique $a_\infty \in \mathscr{L}_\infty$ such that $a_\infty < a$ and a_∞ contains all elements of $\mathscr{L}[0,a]$ that lie at infinity; a_∞ is the unique element at infinity contained in a with dimension equal to $\dim(a) - 1$. In fact, let $K \subset J$ be a finite set with $a < O \vee u(K)$; as $a \nless u(K)$, $\dim(a \vee u(K)) > \dim(u(K))$ so that

$$a \vee u(K) = O \vee u(K);$$

if $a_\infty = a \wedge u(K)$, then $\dim(a_\infty) = \dim(a) - 1$, from which the characterizations of a_∞ follow immediately. The uniqueness of a_∞ is obvious and a_∞ is called the *hyperplane at inifinity in a*. In particular, if P is a point not in $\mathscr{L}_\infty$ and $K \subset J$ a finite set, then for $a = P \vee u(K)$, $u(K)$ is the hyperplane at infinity. Every line not in $\mathscr{L}_\infty$ has a unique point at infinity on it, every plane not in $\mathscr{L}_\infty$ has a unique line at infinity on it, and so on.

Let m_j be the line $O \vee P_j$ $(j \in J)$. For each $j \in J$ we choose a point E_j in m_j distinct from O and P_j. As explained earlier we convert the set of points of m_j different from P_j into a division ring $\mathbf{D}_j$ with O and E_j as its null and unit elements, respectively. The division rings $\mathbf{D}_j$ are of course mutually isomorphic but we can do better and introduce a "compatible" system of isomorphisms. Let $j, k \in J$, $j \neq k$. The lines $E_j \vee E_k$ and $P_j \vee P_k$, being in the plane $m_j \vee m_k$, meet at some point, say P_{jk}; this is the point at infinity of $E_j \vee E_k$, and the perspectivity from P_{jk}, of the line m_k onto the line m_j, fixes O and takes E_k (resp. P_k) to E_j (resp. P_j). It thus induces an isomorphism of $\mathbf{D}_k$ onto $\mathbf{D}_j$, which we denote by

$$(21)\qquad \theta_{jk} : \mathbf{D}_k \xrightarrow{\sim} \mathbf{D}_j, \qquad \theta_{jk} = \pi(P_{jk}; m_k, m_j).$$

We define θ_{jj} to be the identity. Then the "compatibility" of the θ_{jk} may be formulated as the following lemma.

Lemma 2.10. *If $j, k, r \in J$, then*

$$(22)\qquad \theta_{rj} \circ \theta_{jk} = \theta_{rk}.$$

Proof. From (21) it is immediate that $\theta_{kj}=\theta_{jk}^{-1}$. Suppose now that j, k, r are three distinct indices from J, $X \in \mathbf{D}_k$, $Y=\theta_{jk}(X)$, $Z=\theta_{rj}(Y)$. We must prove that $Z=\theta_{rk}(X)$. We may assume $X \neq O$, $X \neq E_k$. Desargue's theorem applied to the triangles XYZ and $E_kE_jE_r$ which are in perspective from O, gives a line t on which the points

$$(X \vee Y) \wedge (E_k \vee E_j) = P_{jk}, \qquad (Y \vee Z) \wedge (E_j \vee E_r) = P_{jr},$$

and

$$(X \vee Z) \wedge (E_k \vee E_r) = Q$$

lie. Obviously t is at infinity and so $Q = P_{rk}$, proving that $Z=\theta_{rk}(X)$.

From (22) we see that there is a division ring $\mathbf{D}$ and isomorphisms

$$\phi_j : \mathbf{D} \xrightarrow{\sim} \mathbf{D}_j \tag{23}$$

such that the diagrams

$$\begin{tikzcd} \mathbf{D} \arrow[rr, "\overset{\varphi_i}{\sim}"] \arrow[dr, "\wr", "\varphi_k"'] & & \mathbf{D}_j \\ & \mathbf{D}_k \arrow[ur, "\theta_{jk}", "\sim"'] & \end{tikzcd} \tag{24}$$

are commutative. We write O and E for the null and unit elements of $\mathbf{D}$.

Let $\mathfrak{P}$ be the set of points of $\mathscr{L}$ which do not lie at infinity. Let us denote by $\mathbf{W}$ the vector space of all functions with finite supports on J with values in $\mathbf{D}$. $\mathbf{W}$ is a left vector space over $\mathbf{D}$ in the usual fashion. Suppose now $P \in \mathfrak{P}$. We can choose a finite set $K \subset J$ such that $P < O \vee u(K)$. For any $j \in J$, $u(K-\{j\})$ is the hyperplane at infinity on $P \vee u(K-\{j\})$ and so the latter does not contain P_j; thus

$$M_{jK}{}^P = m_j \wedge (P \vee u(K-\{j\}))$$

is an element of $\mathbf{D}_j$. If $K_1 \subset K$ is such that $P < O \vee u(K_1)$, it is obvious that $M_{j,K}{}^P = M_{j,K_1}{}^P$; if $L \subset K$ is finite with $P < O \vee u(L)$, then $P < O \vee u(K \cap L)$ by (20), and so $M_{j,K}{}^P = M_{j,L}{}^P = M_{j,K\cap L}^P$. We thus write, for any $j \in J$,

$$M_j{}^P = m_j \wedge (P \vee u(K-\{j\})) \qquad (P < O \vee u(K)). \tag{25}$$

Transferring $M_j{}^P$ to $\mathbf{D}$ now gives the coordinates of P, i.e., let $\mathbf{f}_P$ be the function from J to $\mathbf{D}$ defined by

$$\mathbf{f}_P(j) = \varphi_j^{-1}(M_j{}^P) \qquad (j \in J). \tag{26}$$

Lemma 2.11. *For any $P \in \mathfrak{P}$, $\mathbf{f}_P$ is in $\mathbf{W}$; $\mathbf{f}_O = \mathbf{0}$, the origin of $\mathbf{W}$; and $P \to \mathbf{f}_P$ is a bijection of $\mathfrak{P}$ with $\mathbf{W}$.*

Proof. If $P < O \vee u(K)$ and $j \notin K$, $M_j{}^P = m_j \wedge (P \vee u(K)) = O$, showing that $\mathbf{f}_P \in \mathbf{W}$. It is also obvious that $\mathbf{f}_O = \mathbf{0}$. Let $\mathbf{g} \in \mathbf{W}$, $\mathbf{g} \neq \mathbf{0}$, and let $J' = \{j_1, \dots, j_s\}$ be the support of $\mathbf{g}$. Let $Q_r = \varphi_{j_r}(\mathbf{g}(j_r))$, $a_r = Q_r \vee u(J'-\{j_r\})$, $P = \bigwedge_{1 \leqslant r \leqslant s} a_r$. A simple dimension calculation shows that P is a point. If P were to lie at infinity, $P < u(J'-\{j_r\})$ for all r, which is impossible since

$$\bigwedge_{1 \leqslant r \leqslant s} u(J'-\{j_r\}) = 0$$

by (20). So, $P \in \mathfrak{P}$ and it is easy to verify that $\mathbf{f}_P = \mathbf{g}$. If $P' \in \mathfrak{P}$ is also such that $\mathbf{f}_{P'} = \mathbf{g}$, it is clear that $P' < O \vee u(J')$, and hence, for $1 \leq r \leq s$,

$$Q_r = M_{j_r}{}^{P'} < P' \vee u(J' - \{j_r\})$$

so that $P' < a_r$ for all r, showing that $P' = P$.

Remark. For later use we mention here the following easily verified fact: let $P \in \mathfrak{P}$ be different from O; R, the point at infinity on the line $O \vee P$; and $j \in J$ any index such that $P \nless m_j$. Then we have a well-defined point S, necessarily at infinity, where the line $R \vee P_j$ meets $u(K - \{j\})$; and for any $Q \in \mathfrak{P}$, $Q < O \vee P$, $Q \neq O$, the line $Q \vee M_j{}^Q$ passes through S.

The map $P \to \mathbf{f}_P$ describes the "affine" geometry of $\mathfrak{P}$ in terms of the affine linear geometry of $\mathbf{W}$. The next three lemmas make precise this statement, and they constitute the heart of the proof of the main theorem of this section.

Lemma 2.12. *Let $P \in \mathfrak{P}$, $P \neq O$. Then a point Q of $\mathfrak{P}$ lies on the line $O \vee P$ if and only if for some $M \in \mathbf{D}$ one has*

$$\mathbf{f}_Q = M \cdot \mathbf{f}_P.$$

Proof. Let $K = K_P \subseteq J$ be the smallest finite set $\subseteq J$ such that

$$P < O \vee u(K);$$

to avoid trivialities we assume that K has at least two elements. Let $Q \neq O$ be on the line $O \vee P$, $Q \in \mathfrak{P}$. As $K_Q = K$, $\mathbf{f}_Q(j) \neq O$ for all $j \in K$ and so for each $j \in K$ there is a unique nonzero $M_j \in \mathbf{D}$ such that $M_j{}^Q = M_j \cdot M_j{}^P$. We need to show that for $j, k \in K$, $j \neq k$, $\theta_{jk}(M_k) = M_j$. If for some $j \in K$, $M_j{}^Q = M_j{}^P$ then our remark above shows that $P = Q$ and so we may suppose that $M_j{}^Q \neq M_j{}^P$, $j \in K$. Fix $j, k \in K$, $j \neq k$. The triangles $PM_j{}^P M_k{}^P$ and $QM_j{}^Q M_k{}^Q$ are in perspective from O and so have an axis of perspectivity. This is a line at infinity (cf. the remark above) and so

$$(M_j{}^P \vee M_k{}^P) \wedge (M_j{}^Q \vee M_k{}^Q) = T$$

is a point at infinity, which lies on the line $P_j \vee P_k$ because all four points are in the plane $a = O \vee P_j \vee P_k$. If π denotes the perspectivity, in the plane $a = O \vee P_j \vee P_k$, of m_j to m_k with center T, we can use the definition of multiplication in $\mathbf{D}_j$ to conclude that if $N_j = (\theta_{jk} \circ \pi)(E_j)$, then for any $U \in \mathbf{D}_j$, $(\theta_{jk} \circ \pi)(U) = U \cdot N_j$. Thus $\theta_{jk}(M_k{}^Q) = M_j{}^Q \cdot N_j$ and $\theta_{jk}(M_k{}^P) = M_j{}^P \cdot N_j$, proving that $\theta_{jk}(M_k) = M_j$.

For the converse, suppose $Q \neq O$ is in $\mathfrak{P}$ with $\mathbf{f}_Q = M \cdot \mathbf{f}_P$. Assume as before, to avoid trivialities, that K has at least two elements. Choose a $j_0 \in K$ and let Q' be the point on the line $O \vee P$ such that the lines $Q' \vee M_{j_0}{}^Q$ and $P \vee M_{j_0}{}^P$ meet at infinity. Obviously, $Q' \in \mathfrak{P}$, $Q' \neq O$, and $M_{j_0}{}^{Q'} = M_{j_0}{}^Q$. Since

$\mathbf{f}_{Q'} = M' \cdot \mathbf{f}_P$ for some $M' \in \mathbf{D}$, we have $\varphi_{j_0}(M') = \varphi_{j_0}(M)$ so that $M' = M$, i.e., $Q = Q'$.

Lemma 2.13 (*"parallelogram law of addition"*). *Let $O' \in \mathfrak{P}$, $O' \neq O$, and let t' be a line through O' not containing O. Let U, R be the points at infinity on the lines $O \vee O'$ and t', respectively, and let t be the line $O \vee R$. If $P \in \mathfrak{P}$ lies on t and*

$$P' = (U \vee P) \wedge t', \tag{27}$$

then

$$\mathbf{f}_{P'} = \mathbf{f}_P + \mathbf{f}_{O'}.$$

Conversely, suppose $P' \in \mathfrak{P}$ lies on t'; then $\mathbf{f}_{P'} - \mathbf{f}_{O'} = \mathbf{f}_P$ for some P on t, and P and P' are related by (27).

Proof. For the first part we must show that $M_j{}^{P'} = M_j{}^P + M_j{}^{Q'}$ for all $j \in K$ where $K \subseteq J$ is a finite set such that O', P, P' are all $< O \vee u(K)$. If $U < O \vee u(K - \{j\})$, it is easy to see that $M_j{}^{P'} = M_j{}^P$ and $M_j{}^{O'} = O$; if $R < O \vee u(K - \{j\})$, it is likewise easy to check that $M_j{}^P = O$ and $M_j{}^{P'} = M_j{}^{Q'}$. So assume that neither U nor R lies in $O \vee u(K - \{j\})$. Then $t' \wedge (O \vee u(K - \{j\}))$ is a point, say O''; $O'' \neq O$, and $O'' \in \mathfrak{P}$. Let U' be the point at infinity on $O \vee O''$, and P'' and O_1, the respective points on $O \vee R$ where $U' \vee P'$ and $U' \vee O'$ meet it (see Fig. 3).

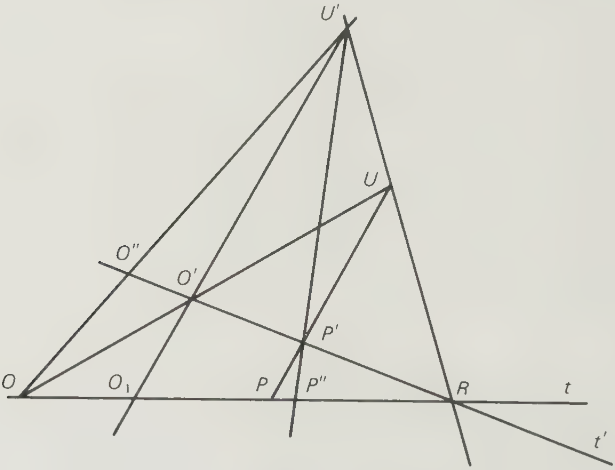

FIG. 3

As $O'' < O \vee u(K - \{j\})$ we have $U' < u(K - \{j\})$ and so

$$M_j{}^{P'} = M_j{}^{P''}, \qquad M_j{}^{O'} = M_j{}^{O_1}.$$

It is clear from Fig. 3 that with respect to addition on the line t (minus R), $P'' = P + O_1$; so, if $t = m_j$, $M_j{}^{P''} = M_j{}^P + M_j{}^{O_1}$. If $t \neq m_j$, let T be the common point at infinity on the lines $O_1\vee M_j^{O_1}$, $P\vee M_j^{P}$, $P''\vee M_j^{P''}$. Consider the plane $t\vee m_j$, and in it, the perspectivity from t to m_j with center T. As this takes O_1, P, P'' to $M_j^{O_1}$, M_j^{P}, and $M_j^{P''}$, respectively, we get

$$M_j^{P''} = M_j^{P} + M_j^{O_1}.$$

For the converse let $\mathbf{g}=\mathbf{f}_{P'}-\mathbf{f}_{O'}$. Define $P=t\wedge(U\vee P')$. Then P and P' are related by (27) and so $\mathbf{g}=\mathbf{f}_{P}$.

Lemma 2.14. *Let P_1', P_2', P_3' be three distinct points in $\mathfrak{P}$. Then they are collinear if and only if for some $A\neq O$ in $\mathbf{D}$ we have*

$$\mathbf{f}_{P_3'} = A\cdot\mathbf{f}_{P_1'}+(E-A)\cdot\mathbf{f}_{P_2'}. \tag{28}$$

Proof. Assume first that the P_i' are on a line t'. Then (28) follows from lemma 2.12 if t' contains O. If $O\nless t'$, we can apply lemma 2.13 with $O'=P_2'$. If P_1 and P_3 are the points on the line $O\vee R$ corresponding to P_1' and P_3', we have

$$\mathbf{f}_{P_3'} = \mathbf{f}_{P_2'}+\mathbf{f}_{P_3}, \qquad \mathbf{f}_{P_1'} = \mathbf{f}_{P_2'}+\mathbf{f}_{P_1},$$

while lemma 2.12 gives a nonzero $A\in\mathbf{D}$ with $\mathbf{f}_{P_3}=A\cdot\mathbf{f}_{P_1}$. This implies (28). For the converse, assume (28) and let t' be the line $P_1'\vee P_2'$. We may suppose that $O\nless t'$, as the case $O<t'$ is trivial. Write $O'=P_2'$ and define P_1 corresponding to P_1'. If P_3 is the point on the line $O\vee P_1$ such that $\mathbf{f}_{P_3}=A\cdot\mathbf{f}_{P_1}$, then $\mathbf{f}_{P_3'}=\mathbf{f}_{P_2'}+\mathbf{f}_{P_3}$. Lemma 2.13 now proves that $P_3'<t'$.

We are now in a position to introduce *homogeneous* coordinates for all points of $\mathscr{L}$. Let I be an index set containing J and one extra element which we denote by ∞. Let $\mathbf{V}$ be the left vector space of all $\mathbf{D}$-valued functions on I with finite supports, and $\mathscr{L}'$, the generalized geometry of all finite dimensional subspaces of $\mathbf{V}$. For any $P\in\mathfrak{P}$ we define $\mathbf{g}_P\in\mathbf{V}$ and $\gamma(P)\in\mathscr{L}'$ by

$$\mathbf{g}_P(j) = \begin{cases}\mathbf{f}_P(j), & j\in J,\\ E, & j=\infty,\end{cases} \tag{29}$$

and

$$\gamma(P) = \mathbf{D}\cdot\mathbf{g}_P. \tag{30}$$

Suppose P is a point at infinity on $\mathscr{L}$. We choose a point Q on the line $O\vee P$ such that $Q\neq P$, $Q\neq O$ and define $\mathbf{g}_{P,Q}\in\mathbf{V}$, $\gamma(P)\in\mathscr{L}'$ by

$$\mathbf{g}_{P,Q}(j) = \begin{cases}\mathbf{f}_Q(j), & j\in J,\\ O, & j=\infty,\end{cases} \tag{31}$$

and

$$\gamma(P) = \mathbf{D}\cdot\mathbf{g}_{P,Q}. \tag{32}$$

Lemma 2.12 shows that $\gamma(P)$ in (32) is independent of the choice of Q and so is well defined.

Lemma 2.15. *The map $P \to \gamma(P)$ is a bijection, of the set of all points of $\mathscr{L}$ with the set of all points of $\mathscr{L}'$, which preserves collinearity.*

Proof. The bijective nature of the map $P \to \gamma(P)$ is a straightforward consequence of the definitions and we sketch the argument for proving that γ preserves collinearity. Fix three points of $\mathscr{L}$, say P_i $(i=1,2,3)$; we must show that they are collinear in $\mathscr{L}$ if and only if $\gamma(P_i)$ are collinear in $\mathscr{L}'$.

Case 1. $P_i \in \mathfrak{P}$ for all i. In this case lemma 2.14 gives what we want.

Case 2. Exactly one of the P_i, say P_3, lies at infinity. Let Q be a point on $O \vee P_3$, $Q \neq O$, $Q \neq P_3$. It is trivial that $\gamma(P_1)$, $\gamma(P_2)$, and $\gamma(P_3)$ are collinear if and only if for some $M \in \mathbf{D}$, $\mathbf{f}_Q = M \cdot (\mathbf{f}_{P_2} - \mathbf{f}_{P_1})$. Lemmas 2.12 and 2.13 show that this is precisely the condition for Q to lie on the line joining O to the point at infinity on $P_1 \vee P_2$, i.e., for P_3 to lie on $P_1 \vee P_2$.

Case 3. P_2 and P_3 are both at infinity. Suppose P_1, P_2, P_3 are collinear, so that P_1 is also at infinity. Chooce $Q_i \in \mathfrak{P}$, $Q_i \neq O$ $(i = 1,2)$ such that: (i) $Q_i < O \vee P_i$; (ii) Q_1, Q_2, and P_3 are collinear. Choose $Q_3 \in \mathfrak{P}$, $Q_3 \neq O$ on $O \vee P_3$. By lemmas 2.12 and 2.13 $\mathbf{f}_{Q_3} = M \cdot (\mathbf{f}_{Q_2} - \mathbf{f}_{Q_1})$ for some $M \in \mathbf{D}$. This shows that $\gamma(P_1)$, $\gamma(P_2)$, and $\gamma(P_3)$ are collinear. Conversely, if $\gamma(P_1)$, $\gamma(P_2)$, $\gamma(P_3)$ are collinear, choose $Q_i \in \mathfrak{P}$, $Q_i \neq O$ in $O \vee P_i$, $i = 1,2,3$. Then for some M, $N \neq O$ in $\mathbf{D}$,

$$\mathbf{f}_{Q_3} = M \cdot \mathbf{f}_{Q_1} + N \cdot \mathbf{f}_{Q_2}.$$

By lemma 2.12 we can find points $Q_j' \neq O$, such that

$$\mathbf{f}_{Q_1'} = -M \cdot \mathbf{f}_{Q_1}, \qquad \mathbf{f}_{Q_2'} = N \cdot \mathbf{f}_{Q_2}.$$

Then

$$\mathbf{f}_{Q_3'} = \mathbf{f}_{Q_2'} - \mathbf{f}_{Q_1'}.$$

If X is the point at infinity on $Q_1' \vee Q_2'$, lemma 2.13 and this relation tell us that Q_3 is on $O \vee X$. So $X = P_3$, i.e., the P_i $(i=1,2,3)$ are collinear. This proves the lemma.

These lemmas imply the following result which is the fundamental theorem on coordinatization of geometries.

Theorem 2.16. *Let $\mathscr{L}$ be a generalized geometry with* $\dim(\mathscr{L}) \geqslant 4$. *Then there exists a division ring* $\mathbf{D}$ *and a left vector space* $\mathbf{V}$ *over* $\mathbf{D}$ *such that $\mathscr{L}$ is isomorphic to the generalized geometry of all finite dimensional subspaces of* $\mathbf{V}$. $\mathbf{V}$ *is finite dimensional if and only if $\mathscr{L}$ is a geometry.*

NOTES ON CHAPTER II

1. For more detail on quaternionic geometry see Chapter I of Chevalley [1]. The identification of $\mathbf{H}$ with $\mathbf{C}^2$ in Section 2 gives rise to a bijection

$\mathbf{H}^n \xrightarrow{\sim} \mathbf{C}^{2n}$ which is $\mathbf{C}$-linear and associates $\mathbf{q}=(q_1, \ldots, q_n)$ to $x=(x_1, \ldots, x_{2n})$ where $\mathbf{q}_k = x_k\mathbf{1} + x_{n+k}\mathbf{j}_2$. We may regard $\mathbf{H}$ as a Hilbert space with scalar product $(\mathbf{q},\mathbf{r}) = \sum_{1\leqslant k\leqslant n} \mathbf{q}_k\mathbf{r}_k^*$ where $\mathbf{q}=(\mathbf{q}_1, \ldots, \mathbf{q}_n)$ and $\mathbf{r}=(\mathbf{r}_1, \ldots, \mathbf{r}_n)$. The group of linear automorphisms of $\mathbf{H}^n$ that preserves $(\,.\,,\,.\,)$ goes over to the subgroup of $U(2n,\mathbf{C})$ that preserves the symplectic form

$$\langle x,y\rangle = \sum_{1\leqslant k\leqslant n} (x_k y_{n+k} - y_k x_{n+k}),$$

where $x=(x_1, \ldots, x_{2n})$ and $y=(y_1, \ldots, y_{2n})$.

2. The definitions in Section 5 are classical. For any line t in a Desarguesian plane π let G be the set of all projectivities of t with itself which are compositions of *two* perspectivities: from t to a second line s followed by a perspectivity from s to t. The line s, as well as the centers of perspectivity, are arbitrary. Assume that $s \neq t$ and that the two centers, X, Y are distinct. Then the projectivity has one or two fixed points according as the line $X \vee Y$ passes through $s \wedge t$ or not. In the former case the projectivity is called *special* and the point W, which is on s, t and $X \vee Y$, is the unique fixed point and is called its *canonical* fixed point. If $O = s\wedge t$ and $W = (X \vee Y)\wedge t$ are distinct, they are the two fixed points and are respectively called the *first* and *second canonical* fixed points of the projectivity. For any W on t (resp. distinct O, W, on t), $G_{s,W}$ (resp. $G_{g,O,W}$) is the set consisting of the identity and the special (resp. general) projectivities of t with itself, with W as canonical fixed point (resp. O and W as the first and second canonical fixed points).

The key fact, which follows from the Desarguesian nature of π, is that if $\sigma \in G_{s,W}$ (resp. $\sigma \in G_{g,O,W}$), σ is determined by its action on one point of t other than W (resp. its action on one point of t other than O and W), independently of the choices of s, X, Y. It is immediate from this that the operators a_A and m_A are well defined. The possibility of using different s, X, Y to define the same projectivity then allows one to prove that $G_{s,W}$ and $G_{g,O,W}$ are groups, the former being abelian. This will prove that $+$ converts $\mathbf{D}$ into an abelian group with O as its null element, and $\cdot$ converts $\mathbf{D}\backslash 0$ into a group with E as unit element. Right multiplication is of course in $G_{g,O,W}$; left multiplication is also a projectivity of t with itself fixing O and W, which however is a composition of *three* perspectivities. Distributivity laws are special cases of the more general result that if α is any projectivity of t fixing O and W, α is an automorphism of the additive group of $\mathbf{D}$. This is an immediate consequence of the extension principle stating that any projectivity from one line of π to another can be extended to an automorphism of π. This principle, in the case of a perspectivity, is an easy application of the following lemma:

Lemma. *Let C be a point of π; t, t' distinct lines through C; and A, A' distinct points on t', different from C. Then there is a unique automorphism α*

of the plane π such that: (i) *α fixes each point of t*; (ii) $\alpha(A)=A'$; (iii) *for any point X not on t, $\alpha(X)$ is on the line joining C to X.*

Proof. Consider first a line u through C distinct from t and two points X, X' on u, distinct and different from C. Let $P(u)$ be the set of points of the plane not on u. We define $\theta_{X,X'}$ to be the map $P(u) \rightarrow P(u)$ such that for any B in $P(u)$, $B'=\theta_{X,X'}(B)$ is the point where the line $K \vee X'$ meets the line $C \vee B$, K being the point where the line $B \vee X$ meets t (draw a figure). It is easy to see that $\theta_{X,X'}$ is bijective and its fixed points are precisely those on t (except for C). If v is the line $B \vee B'$ for some B in $P(u)$, $\theta_{B,B'}$ is well defined on $P(v)$, and a suitable application of Desargues' theorem shows that $\theta_{B,B'}=\theta_{X,X'}$ on $P(u) \cap P(v)$. We now define α as follows. For any point $P<t$, $\alpha(P)=P$; if $P \nless t$ and $P \nless t'$, $\alpha(P)=\theta_{A,A'}(P)$; if $P<t'$, we take X outside t and t' and define $\alpha(P)=\theta_{X,X'}(P)$ where $X'=\theta_{A,A'}(X)$. It is then not difficult to show that α is well defined, defines an automorphism of π with the required properties, and is the only such.

3. As a final reference on the foundational aspects of projective geometry we mention the first volume of H. F. Baker's great six-volume classical treatise (H. F. Baker, *Principles of Geometry*; originally published by the Cambridge University Press, and republished in 1968 by Frederick Ungar Publishing Co, New York).

4. To von Neumann we owe the most profound and far reaching of the coordinatization theorems for complemented lattices. Motivated by the examples of projection lattices arising from operator algebras which are factors of type II_1 he carried out a penetrating analysis of complemented modular lattices that are not of finite rank and hence do not have points ("pointless geometries"!). Before doing this the classical theorem of coordinatization had to be reformulated so that it made no reference to points; von Neumann viewed it as the establishing of an isomorphism of the geometry with the lattice of right ideals of the ring of $N \times N$ matrices over some division ring $\mathbf{D}$. In the general case he proved that if $\mathscr{L}$ is a complemented modular lattice with the property that, for some integer $N \geqslant 4$, 1 is the sum of independent elements $(a_i)_{1 \leqslant i \leqslant N}$ any two of which are "in perspective" (i.e., have a common complement), then $\mathscr{L}$ is isomorphic to the lattice of *principal* right ideals of a so-called *regular* ring R, which is determined uniquely up to isomorphism. Here, regularity means that for any $a \in R$, there is $x \in R$ such that $axa=a$; he actually constructed R as the ring of $N \times N$ matrices over another regular ring S. It must be remarked that, in general, the partially ordered set of right ideals of a ring is not complemented, and that if a right ideal A possesses a complementary right ideal B (i.e., $A+B=R$, $A \cap B=(0)$), then A and B are both principal; and further, that the regularity of the ring is necessary and sufficient for the set

of principal right ideals to form a complemented partially ordered set; and finally, that in this case the set of principal right ideals is a complemented modular lattice (so that principal = finitely generated). Of course in all of this "right" can be replaced by "left". The lattice will be irreducible if and only if R is indecomposable in the sense that R does not admit any decomposition of the form $R = R_1 \oplus R_2$ where $R_i \neq 0$ are two-sided ideals; and this is equivalent to requiring that the center of R is a field. Finally, if a ring R has the property that its principal right (or left) ideals satisfy the ascending and descending chain conditions, then R is regular if and only if it is semisimple, thus tying in the classical coordinatization theory with the Wedderburn theory.

It was also established by von Neumann that dualities of $\mathscr{L}$ correspond naturally to anti-automorphisms of R, and that orthocomplementations correspond to involutive anti-automorphisms $x \rightarrow x^*$ which are *definite* in the sense that $x^*x = 0 \Leftrightarrow x = 0$.

Rings of bounded operators in a complex Hilbert space are regular only when they are finite dimensional; however, the ring of not necessarily bounded operators *affiliated* (in the sense of Murray and von Neumann) to a type II_1 factor is a regular ring, and is in fact the regular ring whose principal right ideals form the lattice isomorphic to the lattice of projections of the factor.

Occupying a special place among the complemented modular lattices are the continuous geometries also discovered by von Neumann. He characterized completely the associated regular rings. The lattices of projections of type II_1 factors are orthocomplemented continuous geometries. Roughly speaking the axioms for an irreducible complemented modular lattice to be a continuous geometry require the completeness of the lattice and the continuity of the lattice operations, and lead to the existence of a dimension function with values in the unit interval [0,1].

Let $\mathbf{D}$ be a division ring, V_k and V_{mk} (m an integer $\geqslant 2$) be left vector spaces of dimensions k and mk, respectively, over $\mathbf{D}$. We then have a map $M \rightarrow M \oplus \ldots \oplus M$ (m summands) of $\mathscr{L}(V_k, \mathbf{D})$ into $\mathscr{L}(V_{mk}, \mathbf{D})$. Using a succession of such mappings (with changing "ratios" m), von Neumann constructed continuous geometries associated with $\mathbf{D}$ which, when $\mathbf{D} = \mathbf{C}$, were not isomorphic to the projection lattices of type II_1 factors.

These scattered remarks should convey a glimpse of the beautiful generalization of classical projective geometry that von Neumann erected. The reader who wants to get a sharper perspective on these things should begin with a study of von Neumann's book (*Continuous Geometry*, Princeton University Press, Princeton, New Jersey, 1960, which has annotations by I. Halperin that illuminate many aspects of the theory and contain references to further literature), and his papers on the subject that appear in volume IV of his *Collected Works* (6 volumes, Pergamon, 1961).

CHAPTER III

THE LOGIC OF A QUANTUM MECHANICAL SYSTEM

1. LOGICS

Let $\mathfrak{S}$ be any quantum mechanical system. We have seen that one can associate with $\mathfrak{S}$ the partially ordered set $\mathscr{L}(\mathfrak{S})$ of all experimentally verifiable propositions concerning $\mathfrak{S}$. The partial ordering is that induced by the implication relation. Moreover, the map, which associates with any element of $\mathscr{L}(\mathfrak{S})$ its negation, behaves very much like an orthocomplementation. This leads one to introduce axiomatically a class of partially ordered sets and to study their properties. We call these systems *logics*. The basic assumption of modern quantum theory can then be described (if we anticipate some terminology) by saying that $\mathscr{L}(\mathfrak{S})$ is a standard logic.

Let $\mathscr{L}$ be a lattice under a partial ordering $<$. By an *orthocomplementation* of $\mathscr{L}$ we mean a mapping

$$(\perp) : a \to a^{\perp}$$

of $\mathscr{L}$ into itself such that

$$(1) \qquad \begin{array}{rl} \text{(i)} & \perp \text{ is one-one and maps } \mathscr{L} \text{ onto itself,} \\ \text{(ii)} & a < b \text{ implies } b^{\perp} < a^{\perp}, \\ \text{(iii)} & a^{\perp\perp} = a \quad \text{for all } a, \\ \text{(iv)} & a \wedge a^{\perp} = 0 \quad \text{for all } a, \\ \text{(v)} & a \vee a^{\perp} = 1 \quad \text{for all } a. \end{array}$$

We note that the relations $0^{\perp}=1$ and $1^{\perp}=0$ are easy consequences of (ii). Moreover it follows easily from (iii) that $\perp$ is one-one and onto, so that (iii) $\Rightarrow$ (i). Finally, we observe that (ii), (iii), and (iv) imply (v). In fact, let $a \in \mathscr{L}$ and $b = a \vee a^{\perp}$. Since both a and $a^{\perp}$ are $<b$ and $a^{\perp\perp}=a$, (ii) implies that $b^{\perp}<a$ and $b^{\perp}<a^{\perp}$. Hence $b^{\perp}<a \wedge a^{\perp}=0$ or $b=b^{\perp\perp}=1$. A lattice $\mathscr{L}$, with an orthocomplementation $\perp(a \to a^{\perp})$, is said to be a *logic* if

(2)
- (i) for any countably infinite sequence $a_1, a_2, \cdots$ of elements of $\mathscr{L}$, $\bigvee_n a_n$ and $\bigwedge_n a_n$ exist in $\mathscr{L}$,
- (ii) if $a_1, a_2 \in \mathscr{L}$ and $a_1 < a_2$, there exists an element $b \in \mathscr{L}$ such that $b < a_1^{\perp}$ and $b \vee a_1 = a_2$.

Before proceeding further we make the observation that if $a_1 < a_2$, the element b whose existence is postulated in (ii) of (2) is unique and is in fact equal to $a_1^{\perp} \wedge a_2$. To see this, let us write $b_1 = a_1^{\perp} \wedge a_2$. Since $b < a_1^{\perp}$ and $b < a_2$, we have $b < b_1$. Therefore, using (ii) of (2) again, we find a $d \in \mathscr{L}$ such that $d < b^{\perp}$ and $d \vee b = b_1$. Since $d < b^{\perp}$, $b < d^{\perp}$. Further, $d < b_1 < a_1^{\perp}$ so that $a_1 < d^{\perp}$. Therefore, $a_1 \vee b = a_2 < d^{\perp}$ which implies that $d < a_2^{\perp}$. Moreover, $d < b_1 < a_2$ also. Consequently, $d < a_2 \wedge a_2^{\perp} = 0$ which means that $b_1 = b \vee d = b$.

The element $b = a_1^{\perp} \wedge a_2$ of (ii) of (2) will often be written as $a_2 - a_1$. For any $a \in \mathscr{L}$, we call $a^{\perp}$ the *orthogonal complement* of a in $\mathscr{L}$. If $a < b^{\perp}$, then $b < a^{\perp}$; we shall then say that a and b are *orthogonal* and write $a \perp b$. Suppose $a, b \in \mathscr{L}$ are two elements such that $a < b$, $a \neq b$. As in Chapter II, we write

$$\mathscr{L}[a,b] = \{c : c \in \mathscr{L}, \quad a < c < b\}. \tag{3}$$

Then under the partial ordering inherited from $\mathscr{L}$, $\mathscr{L}[0,b]$ becomes a lattice, in which countable unions and intersections exist and whose zero element is 0 and unit element is b. If we define, for any x in $\mathscr{L}[0,b]$, its orthogonal complement x' by $x' = x^{\perp} \wedge b$, then it can be shown that $\mathscr{L}[0,b]$ equipped with the orthocomplementation $x \rightarrow x'$ is a logic (cf. corollary 3.3).

The central assumption that one makes in any quantum mechanical application is that *the set of experimentally verifiable propositions is a logic*. The only thing that may be open to serious question in this is assumption (i) of (2) which forces any two elements of $\mathscr{L}$ to have a lattice sum; the others may be regarded as technical necessities. We can offer no really convincing phenomenological argument to support this (cf. Birkhoff-von Neumann [1]). If we omit the assumption that $\mathscr{L}$ is a lattice the axioms become so weak that it is very difficult to avoid pathology in any mathematical discussion. We point out also that every calculation in quantum mechanics is based on assumptions which not only imply that the set of experimentally verifiable propositions is a logic but in fact a very special one.

We shall now make a few general remarks. If $\mathscr{L}_1$ and $\mathscr{L}_2$ are two logics, an *injection* of $\mathscr{L}_1$ into $\mathscr{L}_2$ is a map $f (a \rightarrow f(a))$ of $\mathscr{L}_1$ into $\mathscr{L}_2$ such that (i) f is one-one, $f(0) = 0$, $f(1) = 1$; (ii) if $a_1, a_2, \cdots$ is any at most countable sequence of elements of $\mathscr{L}_1$, $f(\bigvee_n a_n) = \bigvee_n f(a_n)$ and $f(\bigwedge a_n) = \bigwedge f(a_n)$; (iii) $f(a^{\perp}) = f(a)^{\perp}$ for all $a \in \mathscr{L}_1$. An *isomorphism* of $\mathscr{L}_1$ on $\mathscr{L}_2$ is an injection which maps $\mathscr{L}_1$ onto $\mathscr{L}_2$. If $\mathscr{L}_1 = \mathscr{L}_2$, isomorphisms are called *automorphisms*. The set of all automorphisms of a logic $\mathscr{L}$ is a group under composition; we denote it by Aut($\mathscr{L}$). If $\mathscr{L}_1$ and $\mathscr{L}$ are two logics, we say that $\mathscr{L}_1$ is a *sublogic* of $\mathscr{L}$ if (i) $\mathscr{L}_1 \subseteq \mathscr{L}$ and (ii) the identity map of $\mathscr{L}_1$ into $\mathscr{L}$ is an injection.

Any Boolean σ-algebra is a logic provided we define, for any element a, $a^\perp$ to be the complement of a. These logics are of course not very interesting from the point of view of quantum mechanics. To obtain a more typical example, let us consider a Hilbert space $\mathscr{H}$ over the real, complex or quaternionic division rings. We denote by $\mathscr{L}(\mathscr{H})$ the collection of all *closed* linear manifolds of $\mathscr{H}$. If we now define $<$ to mean set inclusion, and $\perp$ to mean the usual operation of orthogonal complementation in $\mathscr{L}(\mathscr{H})$, then it can easily be verified that $\mathscr{L}(\mathscr{H})$ is a logic. The isomorphism class of $\mathscr{L}(\mathscr{H})$ depends only on the field $\mathbf{D}$ of definition of $\mathscr{H}$ and the dimension of $\mathscr{H}$. A logic $\mathscr{L}$ is said to be *standard* if it is isomorphic to the logic $\mathscr{L}(\mathscr{H})$ of a separable infinite dimensional Hilbert space over one of the three division rings $\mathbf{R}$, $\mathbf{C}$, or $\mathbf{H}$. Modern quantum theory works with the assumption that the logic of any atomic system is, if not standard, at least a sublogic of a standard logic. We shall make a deep study of standard logics in Chapter IV. We shall also describe there a number of logics that are sublogics of standard logics, but are not standard themselves.

We shall now proceed to derive a few simple consequences of the axioms defining a logic.

Lemma 3.1. *Let $\mathscr{L}$ be a logic and $a_1, a_2, \ldots$ a sequence of elements of $\mathscr{L}$. If $b \in \mathscr{L}$ and $b \perp a_n$ for all n, then $b \perp \bigvee_n a_n$. Moreover we have the identities:*

$$\Big(\bigvee_n a_n\Big)^\perp = \bigwedge_n a_n^\perp, \qquad \Big(\bigwedge_n a_n\Big)^\perp = \bigvee_n a_n^\perp. \tag{4}$$

Finally, if $a_1 < a_2$, $c < a_2$, and $c \perp a_1$, then $c < a_2 - a_1$.

Proof. Since $b < a_n^\perp$, $a_n < b^\perp$ for all n. Hence $\bigvee_n a_n < b^\perp$ or $b < (\bigvee_n a_n)^\perp$. This proves that $b \perp \bigvee_n a_n$. We now prove the first equation of (4). For any m, $a_m < \bigvee_n a_n$, so that $(\bigvee_n a_n)^\perp < a_m^\perp$. Consequently, $(\bigvee_n a_n)^\perp < \bigwedge_m a_m^\perp$. On the other hand, $\bigwedge_m a_m^\perp < a_n^\perp$ for any n and hence $a_n < (\bigwedge_m a_m^\perp)^\perp$. This leads to the relation $\bigvee_n a_n < (\bigwedge_m a_m^\perp)^\perp$ or $(\bigwedge_m a_m^\perp) < (\bigvee_n a_n)^\perp$. This proves the first relation in (4). The second is obtained from the first by replacing a_n by $a_n^\perp$ for all n and taking the orthogonal complements of the relation obtained. Finally, let $a_1 < a_2$, $c < a_2$, and $c \perp a_1$. Write $a_3 = a_1 \vee c$. $a_3 < a_2$ and hence there is a $d \in \mathscr{L}$ such that $d \perp a_3$ and $a_3 \vee d = a_2$. Since $d \perp a_1$ and $c \perp a_1$, $a_1 \perp (c \vee d)$. Moreover, $a_1 \vee (c \vee d) = a_2$. Hence $c \vee d = a_1^\perp \wedge a_2$, which proves that $c < a_2 - a_1$.

Lemma 3.2. *Let a, b, c be three elements of a logic such that $a \perp b$ and $b < c$. Then*

$$(a \vee b) \wedge c = (a \wedge c) \vee b. \tag{5}$$

Proof. Write $d = (a \wedge c) \vee b$ and $e = (a \vee b) \wedge c$. It is obvious that $d < e$.

Hence, by (ii) of (2), there exists an element g of $\mathscr{L}$ such that $d \perp g$, and $d \vee g = e$. Since $a \perp b$ and the element g satisfies the relations $g < a \vee b$, $g \perp b$, it follows from the last statement of lemma 3.1 that $g < a$. Hence, as $g < e < c$, $g < a \wedge c < d$. However, since $g \perp d$, $g < d^{\perp}$ also. Therefore $g = 0$. This proves that $e = d \vee g = d$.

Corollary 3.3. *Let b be a nonzero element of $\mathscr{L}$. Then $\mathscr{L}[0,b]$ equipped with the orthocomplementation $x \to x'$ defined above is a logic.*

Proof. We must verify (ii), (iii), (iv) of (1) and (ii) of (2). (ii), (iv) of (1) are obvious. We shall now prove (iii) of (1). Let $x < b$. Since $x' \perp x$ and $x' \perp b^{\perp}$, one has $x' \perp x \vee b^{\perp}$, while $x' \vee x \vee b^{\perp} = 1$. Thus, we may conclude that $x'^{\perp} = x \vee b^{\perp}$. But then $x'' = x'^{\perp} \wedge b = (x \vee b^{\perp}) \wedge b = x$ from (5). We come finally to (ii) of (2). Let $x < y < b$. As $\mathscr{L}$ is a logic, there is a $z \in \mathscr{L}$ such that $z \perp x$ and $z \vee x = y$. Since $z < b$, $z = z \wedge b$ so that $z < x'$. This proves (2) in $\mathscr{L}[0,b]$.

Remark. It might be observed that (5) is a special case of the usual modular law. We may therefore regard any logic weakly modular since the modular law (5) is valid, not for all a and b, but only for a, b with $a \perp b$. The standard logics are the only ones which have had any significant measure of success as a framework for the precise formulation and solution of quantum mechanical problems. Since these are *not* modular, it is too restrictive to demand in the definition of a logic that it be modular. This explains why we have rejected modularity as a property of a logic and also why at the same time we have inserted the very much weaker condition (ii) of (2) which leads to (5). On the other hand, if a logic is modular, and separable in a natural sense, then it follows from a general result of Kaplansky [1] that it is a *continuous geometry* in the sense of von Neumann. Whether the logic of any atomic system may be assumed to be an orthocomplemented continuous geometry is still somewhat of an open question. The point is that, unlike a standard logic, a continuous geometry does not have points; this circumstance makes the description of the states of the logic difficult.

2. OBSERVABLES

We shall introduce in this section the concept of an observable associated with a logic. The formal definition is given in the next paragraph, but we shall first try to motivate the definition. Suppose that $\mathfrak{S}$ is a physical system, $\mathscr{L}(\mathfrak{S})$ its logic, and ξ a physical quantity or observable. Then in any experiment which an observer performs, the statements that can be made concerning ξ are of the type which asserts that the value of ξ lies in some set E of real numbers. It is natural and harmless to require that the sets E be Borel. If we denote by $x(E)$ the statement that the value of ξ lies in the Borel set $E \subseteq R^1$, then one has a mapping

$$x : E \rightarrow x(E)$$

of $\mathscr{B}(R^1)$ into $\mathscr{L}(\mathfrak{S})$. We shall regard two observables as "identical" if and only if the corresponding mappings are the same. If f is a real valued Borel function of a real variable, then we mean by $f \circ \xi$ the observable whose value is $f(r)$ whenever ξ takes the value r; to this observable clearly corresponds the mapping $E \rightarrow x(f^{-1}(E))$ (cf. Chapter I, sections 2 and 4).

Motivated by these remarks we introduce the following definitions. Let $\mathscr{L}$ be a logic. An *observable associated with* $\mathscr{L}$ is a mapping

$$x : E \rightarrow x(E)$$

of the σ-algebra $\mathscr{B}(R^1)$ of Borel sets of the real line into $\mathscr{L}$ such that

$$\text{(6)} \quad \begin{array}{ll} \text{(i)} & x(\varnothing) = 0,\ x(R^1) = 1, \\ \text{(ii)} & \text{if } E, F \in \mathscr{B}(R^1) \quad \text{and} \quad E \cap F = \varnothing, \quad x(E) \perp x(F), \\ \text{(iii)} & \text{if } E_1, E_2, \cdots \text{ is a sequence of mutually disjoint Borel sets in } R^1, \\ & \qquad x\left(\bigcup_n E_n\right) = \bigvee_n x(E_n). \end{array}$$

We write $\mathscr{O}$ or $\mathscr{O}(\mathscr{L})$ for the set of all observables associated with $\mathscr{L}$. We remark that the properties (i) through (iii) are natural if we want to interpret $x(E)$ as the statement that the value of the observable ξ lies in E.

If X is any set and $\mathscr{L}$ is an arbitrary σ-algebra of subsets of X, then $\mathscr{L}$ is obviously a logic provided we define $<$ to be set inclusion and $\perp$ to be set complementation. Theorem 1.4 asserts that the set of observables is in canonical one-one correspondence with the algebra of all real valued $\mathscr{L}$-measurable functions on X; if $x(E \rightarrow x(E))$ is any observable, then the corresponding function f on X is the unique $\mathscr{L}$-measurable one such that $f^{-1}(E) = x(E)$ for all $E \in \mathscr{B}(R^1)$.

Suppose that x is an observable and f a real valued Borel function of a real variable. Then

$$f \circ x : E \rightarrow x(f^{-1}(E)) \tag{7}$$

is also an observable.

Lemma 3.4. *Let $\mathscr{L}$ be a logic and x an observable associated with $\mathscr{L}$. Then for any sequence $E_1, E_2, \cdots$ of Borel sets*

$$\begin{aligned} x\left(\bigcup_n E_n\right) &= \bigvee_n x(E_n), \\ x\left(\bigcap_n E_n\right) &= \bigwedge_n x(E_n). \end{aligned} \tag{8}$$

Suppose further that f_1 *and* f_2 *are two real valued Borel functions of a real variable* t. *Let* $f_1 \circ f_2$ *denote the function* $t \to f_1(f_2(t))$. *Then*

$$(f_1 \circ f_2) \circ x = f_1 \circ (f_2 \circ x). \tag{9}$$

Proof. Since $x(E_n) \perp x(R^1 - E_n)$ and $x(E_n) \vee x(R^1 - E_n) = 1$, it follows that $x(R^1 - E_n) = x(E_n)^{\perp}$. Consequently, in view of the equation (4), the second equation in (8) will follow if we prove the first. To prove the first, we begin by observing that if $A \subseteq B$, then $x(A) < x(B)$; in fact,

$$x(B) = x(A) \vee x(B - A).$$

This said, let $E = \bigcup_n E_n$. Then $x(E_n) < x(E)$ for all n so that $\bigvee_n x(E_n) < x(E)$. On the other hand, let $F_1 = E_1$ and for $n > 1$, let $F_n = E_n - \bigcup_{j<n} E_j$. The F_n are then mutually disjoint, $F_n \subseteq E_n$ and $\bigcup_n F_n = E$. Therefore,

$$x(E) = \bigvee_n x(F_n) < \bigvee_n x(E_n).$$

This completes the proof of (8). Equation (9) is trivial.

Let x be an observable associated with $\mathscr{L}$. A real number λ is said to be a *strict* value of x if $x(\{\lambda\}) \neq 0$. x is said to be *discrete* if there exists a countable set $C = \{c_1, c_2, \cdots\}$ of real numbers such that $x(C) = 1$. x is said to be *constant* if there exists a real number a such that $x(\{a\}) = 1$; we shall then write $x = a$. x is said to be *bounded* if there exists a compact set K such that $x(K) = 1$. If we define $\sigma(x)$ by

$$\sigma(x) = \bigcap_{C \text{ closed}, \, x(C) = 1} C, \tag{10}$$

then $\sigma(x)$ is a closed set called the *spectrum* of x. Since the topology of R^1 satisfies the second countability axiom, there exists a *sequence* of closed sets $C_1, C_2, \cdots$ such that $x(C_n) = 1$ for all n and $\sigma(x) = \bigcap_n C_n$. Since $x(R^1 - C_n) = 0$ for all n,

$$\begin{aligned} x(R^1 - \sigma(x)) &= x\Big(\bigcup_n (R^1 - C_n)\Big) \\ &= \bigvee_n x(R^1 - C_n) \\ &= 0, \end{aligned}$$

which proves that $x(\sigma(x)) = 1$. $\sigma(x)$ is thus the *smallest closed set* C such that $x(C) = 1$. The numbers $\lambda \in \sigma(x)$ are called the *spectral values* of x. A strict value is a spectral value but the converse is not true in general. x is bounded if and only if $\sigma(x)$ is a compact set. $\lambda \in \sigma(x)$ if and only if for any open set U containing λ, $x(U) \neq 0$.

Since the defining axioms of a logic are somewhat weak, it is not reasonable to hope for any incisive description of the set of all observables associated with $\mathscr{L}$. In the next chapter we shall examine this question for logics which are associated with Hilbert spaces. For the moment we confine ourselves to a remark on discrete observables. Suppose $\mathscr{L}$ is an arbitrary logic. Let x be a discrete observable and $\{c_1, c_2, \cdots\}$ the set of its strict values. Then $x(\{c_i\})$ and $x(\{c_j\})$ are orthogonal whenever $i \neq j$ and

$$\bigvee_i x(\{c_i\}) = 1.$$

Conversely, let $\{c_i\}_{i \in D}$ be an at most countably infinite set of distinct real numbers and $\{a_i\}_{i \in D}$ a family of mutually orthogonal elements of $\mathscr{L}$ (with the same indexing set D) such that

$$\bigvee_{i \in D} a_i = 1.$$

Then there exists a unique discrete observable x associated with $\mathscr{L}$ such that $x(\{c_i\}) = a_i$ for all $i \in D$. In fact, we have only to define, for $E \in \mathscr{B}(R^1)$,

$$x(E) = \bigvee_{i : c_i \in E} a_i.$$

It is easy to verify that x is an observable. It is clear that $x(\{c_i\}) = a_i$ for all $i \in D$ and that x is discrete. The uniqueness of x is obvious.

Even though many interesting and important observables are discrete, there are quantum observables which are not discrete. The construction of observables which are not discrete is, however, quite complicated. We shall indicate how it can be done in the next chapter for logics which are associated with Hilbert spaces. If $\mathscr{L}$ is an orthocomplemented geometry of finite dimension N, then it is easy to see that every observable associated with $\mathscr{L}$ is discrete and has at most N values. Clearly, such logics are not suitable models to represent the experimentally verifiable propositions of complicated atomic systems.

3. STATES

If $\mathfrak{S}$ is a classical mechanical system, the concept of a *state* of $\mathfrak{S}$ is so defined (cf. Chapter I) that if $\mathfrak{S}$ is in a state p, and ξ is any observable, then ξ has a value in p. This leads one to postulate that the states are points of the phase space and observables are real valued functions on it.

In quantum theory this has been rejected as irreconcilable with the available experimental facts regarding the behavior of atomic systems when subjected to refined microscopic observations (cf. von Neumann [1], Heisenberg [1]). The modern approach to atomic physics is based on the principle that *in a given state of the system, an observable has only a probability distribution of values and in general no sharply defined value; and no matter how carefully the state is prepared, there will be some observables whose values are distributed according to probability distributions* *having arbitrarily high variance.* Accordingly, if $\mathfrak{S}$ is in a state σ, one can associate, with each observable x, a probability distribution, say P_x^σ, on the line, which will be interpreted as the *probability distribution of the values of x when the system is in the state σ.* If x is a discrete observable, and $c_1, c_2, \cdots$ are its possible values, then P_x^σ would have mass concentrated on the set $\{c_1, c_2, \cdots\}$; the probability mass which is then assigned by P_x^σ to the point c_n will be the probability that an experiment on $\mathfrak{S}$, *designed to yield an exact value of x*, gives the value c_n, when $\mathfrak{S}$ is in the state σ. If f is any real valued function of a real variable t, then, for the observable $f \circ x$, we would have the equation:

$$\begin{aligned} &\text{Prob}\{f \circ x = a, \text{ when } \mathfrak{S} \text{ is in the state } \sigma\} \\ &= \sum_{n:f(c_n)=a} P_x^\sigma(\{c_n\}). \end{aligned} \tag{11}$$

With these remarks serving as our motivation we proceed to the formal definitions. Let $\mathscr{L}$ be any logic and $\mathscr{O}$ the set of all observables of $\mathscr{L}$. A *state function* of $\mathscr{L}$ is a map

$$P : x \to P_x \qquad (x \in \mathscr{O}) \tag{12}$$

which assigns to each observable $x \in \mathscr{O}$ a probability distribution P_x on $\mathscr{B}(R^1)$ such that for any real valued Borel function f on R^1 and any observable x,

$$P_{f \circ x}(E) = P_x(f^{-1}(E)). \tag{13}$$

Equation (13) has an interesting consequence. We claim first that if 0 is the zero observable, then P_0 is the probability measure whose entire mass is concentrated at the origin. In fact, if g is the function identically zero, then $g \circ 0 = 0$; since $g^{-1}(E) = R^1$ or $\varnothing$ according as $0 \in E$ or $0 \notin E$, we have, by (13), $P_0(E) = P_0(g^{-1}(E) = 1$ or 0 according as $0 \in E$ or $0 \notin E$. We next observe that if x is any observable and E is a Borel set such that $x(E) = 0$, then $P_x(E) = 0$. In fact, if $x(E) = 0$, then, for the function f whose value is 1 on E and 0 on $R^1 - E$, it is obvious that $y = f \circ x$ is the zero observable and hence P_y is the measure concentrated at the origin. As $E = \{t : f(t) = 1\}$, we have $P_x(E) = P_y(\{1\}) = 0$. In other words,

$$P_x(E) = 0 \qquad (E \in \mathscr{B}(R^1), \quad x(E) = 0). \tag{14}$$

(14) implies that if x is discrete, the mass of P_x is concentrated on the set of strict values of x.

In turns out to be possible to construct all the state functions of $\mathscr{L}$ in a particularly simple manner. In order to describe this construction we introduce, with Mackey [1], the notion of a *question*. An observable x

is said to be a *question* if $x(\{0,1\})=1$; x is then discrete. If $a=x(\{1\})$, it is clear that x is the only question such that $x(\{1\})=a$. We shall call it the *question associated with* a and denote it by q_a.

Suppose now that x is an observable associated with the logic $\mathscr{L}$ and E any Borel set $\subseteq R^1$. If χ_E is the function which takes the value 1 on E and 0 on R^1-E, then $\chi_E \circ x$ is the question associated with $x(E)$, i.e.,

$$\chi_E \circ x = q_{x(E)}. \tag{15}$$

In fact, since χ_E takes only the values 0 and 1, it is clear that $\chi_E \circ x$ is discrete and $(\chi_E \circ x)(\{0,1\})=1$. $\chi_E \circ x$ is therefore a question, and, as E is the set where χ_E takes the value 1,

$$(\chi_E \circ x)(\{1\}) = x(E).$$

The description of all state functions can now be given in terms of the concept of a measure on $\mathscr{L}$. A *measure* on $\mathscr{L}$ is a function

$$p : a \to p(a) \qquad (a \in \mathscr{L}) \tag{16}$$

such that

$$(17) \quad \begin{array}{rl} \text{(i)} & p \text{ is real valued and } 0 \leq p(a) \leq 1 \text{ for all } a \in \mathscr{L}, \\ \text{(ii)} & p(0) = 0,\ p(1) = 1, \\ \text{(iii)} & \text{if } a_1, a_2, \cdots \text{ is a sequence of mutually orthogonal elements of} \\ & \mathscr{L} \text{ and } a=\bigvee_n a_n, \text{ then } p(a)=\sum_n p(a_n). \end{array}$$

Notice that if $a_1, a_2 \in \mathscr{L}$ and $a_1 < a_2$, then there exists, by (ii) of (2), $b \in \mathscr{L}$ such that $a_1 \perp b$ and $a_2 = a_1 \vee b$. Therefore, $p(a_1)+p(b)=p(a_2)$. In particular,

$$p(a_1) \leq p(a_2) \qquad (a_1 < a_2). \tag{18}$$

The concept of a measure on $\mathscr{L}$ is a generalization of the well known concept when $\mathscr{L}$ is a Boolean σ-algebra. However $\mathscr{L}$ need not be distributive and hence the structure of a measure on $\mathscr{L}$ is much more complex than that of a measure on a Boolean σ-algebra.

Theorem 3.5. *Let $\mathscr{L}$ be a logic and $\mathscr{O}$ the set of all observables associated with $\mathscr{L}$. Let p be a measure on $\mathscr{L}$. If we define, for any observable $x \in \mathscr{O}$ and any Borel set $E \subseteq R^1$,*

$$P_x^p(E) = p(x(E)), \tag{19}$$

then P_x^p is a probability measure on $\mathscr{B}(R^1)$ and

$$P^p : x \to P_x^p \tag{20}$$

is a state function of $\mathscr{L}$. Conversely, if $P(x \to P_x)$ is an arbitrary state *function of $\mathscr{L}$, there exists one and only one measure p on $\mathscr{L}$ such that $P_x(E)=p(x(E))$ for all x.*

Proof. Suppose that p is a measure on $\mathscr{L}$. If $E_1, E_2, \cdots$ are mutually disjoint Borel sets of R^1 with union E, then $x(E_1), x(E_2), \cdots$ are mutually orthogonal elements of $\mathscr{L}$ with $x(E)=\bigvee_n x(E_n)$; moreover,

$$P_x{}^p(E) = \sum_n P_x{}^p(E_n).$$

Since $P_x{}^p(R^1)=p(1)=1$, we see that $P_x{}^p$ is a probability measure on $\mathscr{B}(R^1)$. If f is a real valued Borel function of a real variable, then

$$\begin{aligned} P^p_{f\circ x}(E) &= p((f\circ x)(E)) \\ &= p(x(f^{-1}(E))) \\ &= P_x{}^p(f^{-1}(E)). \end{aligned}$$

This proves that $x \to P_x{}^p$ is a state function of $\mathscr{L}$.

Conversely, let $P(x \to P_x)$ be a state function of $\mathscr{L}$. If $a \in \mathscr{L}$, P_{q_a} is a probability distribution on the real line and as q_a is a question, (14) ensures that $P_{q_a}(\{0,1\})=1$. Define

$$p(a) = P_{q_a}(\{1\}). \tag{21}$$

$p(a \to p(a))$ is clearly a well-defined, real valued function on $\mathscr{L}$ and $0 \leq p(a) \leq 1$ for all a. We shall now prove that p is a measure on $\mathscr{L}$. Since q_0 is the question with $q_0(\{1\})=0$, it follows from (14) that $P_{q_0}(\{1\})=0$. This proves that $p(0)=0$. Similarly, we prove that $p(1)=1$. Let $\{a_n\}_{n\geq 1}$ be a sequence of mutually orthogonal elements of $\mathscr{L}$ and let $a=\bigvee_n a_n$. Let x be a discrete observable such that $x(\{0\})=a^{\perp}$ and $x(\{n\})=a_n$, $n=1, 2, \cdots$. Let f_n be the function on R^1 whose value is 1 for the point n and 0 otherwise. Then $f_n \circ x$ is the question q_{a_n} and we have, by (21) and (13),

$$p(a_n) = P_x(\{n\}).$$

Since P_x is a probability measure, it follows that

$$\sum_{n=1}^{\infty} p(a_n) = P_x(Z),$$

where Z is the set $\{1, 2, \cdots\}$. If f denotes the function whose value is 1 on Z and 0 on $R^1 - Z$, $f \circ x$ is the question q_a and hence, exactly as before, $p(a)=P_x(Z)$, so that

$$p(a) = \sum_{n=1}^{\infty} p(a_n). \tag{22}$$

p is thus a measure on $\mathscr{L}$. If x is any observable and $E \in \mathscr{B}(R^1)$, we have, on writing g for the function whose value is 1 on E and 0 on $R^1 - E$,

$$\begin{aligned} P_x(E) &= P_{g \circ x}(\{1\}) \\ &= P_{q_{x(E)}}(\{1\}) \\ &= p(x(E)). \end{aligned}$$

Finally, equation (21) implies easily that p is unique. This completes the proof of the theorem.

In view of the theorem we proved just now, it is natural to define a *state* of a logic to be a measure on $\mathscr{L}$. The set of all states of $\mathscr{L}$ will be denoted by $\mathscr{S}$. If $p(a \rightarrow p(a))$ is a state of $\mathscr{L}$ and $x(E \rightarrow x(E))$ any observable, $E \rightarrow p(x(E))$ is a probability measure on $\mathscr{B}(R^1)$. It is called the *probability distribution of* x *in the state* p and is denoted by $P_x{}^p$.

Using the probability distribution $P_x{}^p$, we may define the concepts of *moments* of an observable. The (*mean*) *expected value* $\mathscr{E}(x|p)$ of an observable x in a state p is defined by

$$\mathscr{E}(x|p) = \int_{-\infty}^{\infty} t\,dP_x{}^p(t). \tag{23}$$

The *variance* of x in the state p is defined when $\mathscr{E}(x^2|p) < \infty$; in this case,

$$\operatorname{var}(x|p) = \mathscr{E}(x^2|p) - [\mathscr{E}(x|p)]^2. \tag{24}$$

Clearly $\operatorname{var}(x|p) \geq 0$ and is $=0$ if and only if $P_x{}^p$ is concentrated at the single point $t_0 = \mathscr{E}(x|p)$, and then it is natural to say that, in the state p, x has a *sharply defined value*, namely t_0.

4. PURE STATES. SUPERPOSITION PRINCIPLE

Let $\mathscr{L}$ be a logic and $\mathscr{S}$ the set of all states of $\mathscr{L}$. If $p_1, p_2, \cdots$ is a sequence of elements of $\mathscr{S}$ and $c_1, c_2, \cdots$ is a sequence of constants such that $c_n \geq 0$ for $n = 1, 2, \cdots$ and $c_1 + c_2 + \cdots = 1$, then the function $p(a \rightarrow p(a))$ defined by

$$p(a) = \sum_{n \geq 1} c_n p_n(a) \tag{25}$$

is easily verified to be a state of $\mathscr{L}$. We shall write

$$p = \sum_{n \geq 1} c_n p_n. \tag{26}$$

$\mathscr{S}$ is therefore a *convex set*. If x is any observable, it follows at once from (26) that

$$P_x^{\,p} = \sum_{n\geq 1} c_n P_x^{\,p_n}. \tag{27}$$

The equation (27) may be given the following interpretation. Suppose that one knows that the state of a system is one of $p_1, p_2, \cdots$ with probabilities $c_1, c_2, \cdots$, respectively. Then, for any bounded observable x, the expectation value of x is $\sum_{n\geq 1} c_n \mathscr{E}(x|p_n)$ which is equal to $\mathscr{E}(x|p)$. In other words, assuming that the system is in the state p is, statistically speaking, equivalent to assuming that its state is one of $p_1, p_2, \cdots$ with respective probabilities $c_1, c_2, \cdots$. It is customary to say that the state p is a *classical superposition* of the states $p_n (n \geq 1)$.

In view of the probability interpretation of (26) and (27), interest centers around states for which no representation of the form (26) is possible. A state p is said to be *pure* if the equation

$$p = cp_1 + (1-c)p_2 \qquad (0 \leq c \leq 1, \quad p_1, p_2 \in \mathscr{S}) \tag{28}$$

implies that $p = p_1 = p_2$. We write $\mathscr{P}$ for the set of all pure states of $\mathscr{L}$. $\mathscr{P}$ is the set of *extreme points* of the convex set $\mathscr{S}$.

We may now introduce the notion of superposition. Let $\mathscr{D}$ be a set of states and p_0 an arbitrary state. We say that p_0 is a *superposition of states in* $\mathscr{D}$ if the following property is satisfied:

$$a \in \mathscr{L}, \quad p(a) = 0 \quad \text{for all} \quad p \in \mathscr{D} \Rightarrow p_0(a) = 0. \tag{29}$$

If $p = c_1 p_1 + c_2 p_2 + \cdots$ where $c_1, c_2 \cdots \geq 0$, and $c_1 + c_2 + \cdots = 1$, then p is a superposition of the states in the set $\{p_1, p_2, \cdots\}$. As we shall presently see, if $\mathscr{L}$ is a Boolean σ-algebra, this is, roughly speaking, the only kind of superposition possible, in particular, no pure state can be a superposition of other pure states distinct from it. This is in contradistinction to quantum mechanics, where the structure of $\mathscr{L}$, namely the fact that it is standard, forces the concept of superposition to be nontrivial *even among pure states*.

Theorem 3.6. *Let $\mathscr{L}$ be a Boolean σ-algebra of subsets of a space X such that* (i) *$\mathscr{L}$ is separable* (*cf. Chapter I*) *and* (ii) *$\{a\} \in \mathscr{L}$ for all $a \in X$. For any $a \in X$ let δ_a be the state defined by*

$$\delta_a(A) = \begin{cases} 1 & \text{if} \quad a \in A \\ 0 & \text{if} \quad a \notin A \end{cases} \tag{30}$$

Then $\{\delta_a : a \in X\}$ is precisely the set of all pure states of $\mathscr{L}$. If $\mathscr{D}$ is any set of pure states, and p_0 an arbitrary pure state, p_0 is a superposition of states of $\mathscr{D}$ if and only if $p_0 \in \mathscr{D}$.

Proof. Let $\{A_1, A_2, \cdots\}$ generate $\mathscr{L}$. That δ_a is pure is trivially verified. Suppose that p is a pure state. If for some $A_0 \in \mathscr{L}$, $0 < p(A_0) < 1$, then, on putting

$$p_1(A) = (p(A_0))^{-1} p(A \cap A_0),$$
$$p_2(A) = (1 - p(A_0))^{-1} p(A \cap (X - A_0)),$$

we would obtain the decomposition

$$p = p(A_0) \cdot p_1 + (1 - p(A_0)) \cdot p_2;$$

$p_1 \neq p_2$ as $p_1(A_0) = 1$ and $p_2(A_0) = 0$. This is not possible as p is pure by assumption. Therefore we may conclude for any $A \in \mathscr{L}$, that $p(A) = 0$ or 1. By replacing A_n by $X - A_n$ if necessary, we may assume that $p(A_n) = 1$ for all n, where $\{A_1, A_2, \cdots\}$ generates $\mathscr{L}$. Let

$$B = \bigcap_n A_n.$$

Then $p(B) = 1$. In particular, B is nonempty. We assert that B cannot contain more than a single point. In fact, the collection of all sets $C \in \mathscr{L}$ with the property that either $B \subseteq C$ or $B \cap C = \varnothing$, is a σ-algebra which contains all the A_n. Hence it coincides with $\mathscr{L}$. Therefore, as $\{x\} \in \mathscr{L}$, this is possible only if $B = \{a\}$ for some $a \in X$. But then $p = \delta_a$. Finally, let p_0 be a superposition of the states of $\mathscr{D}$ where each element of $\mathscr{D}$ is pure. If $p_0 = \delta_{a_0}$, but $p_0 \notin \mathscr{D}$, then $p(\{a_0\}) = 0$ for all $p \in \mathscr{D}$ but $p_0(\{a_0\}) = 1$; a contradiction. The theorem is completely proved.

Let $\mathscr{L}$ be an arbitrary logic and $\mathscr{P}$ the set of all pure states of $\mathscr{L}$. For any subset $S \subseteq \mathscr{P}$, we write $\bar{S}$ for the set of all $p \in \mathscr{P}$ such that p is a superposition of the states of S. Let us now consider $\mathscr{M}$, the collection of all $S \subseteq \mathscr{P}$ with the property that $S = \bar{S}$. Under set inclusion, $\mathscr{M}$ becomes a partially ordered set. If $\mathscr{L}$ satisfies the assumptions of theorem 3.6 for example, then $\mathscr{M}$ is the class of all subsets of $\mathscr{P}$ and is therefore a Boolean algebra. However, in the case of more complex $\mathscr{L}$, $\mathscr{M}$ is not a Boolean algebra. In general, the geometric structure of $\mathscr{M}$ seems somewhat hard to determine. For the so-called *standard* logics to be introduced in the next chapter, this can be done and it becomes possible to determine the structure of $\mathscr{M}$ completely. For standard $\mathscr{L}$, $\mathscr{M}$ then becomes *isomorphic* to $\mathscr{L}$. It is this fact which confers on the set of pure states of a standard logic a geometric structure characteristic of quantum theory. In the physical literature (cf. Dirac [1]) this is elevated to a physical principle, the so-called *superposition principle*. From our point of view, the geometry of pure states will appear as a *consequence* of the structure of the logic $\mathscr{L}$.

5. SIMULTANEOUS OBSERVABILITY

One of the features which sets apart quantum mechanics from classical mechanics is the existence of observables whose values cannot always be simultaneously measured. We shall now introduce this idea formally. We begin with the elements of $\mathscr{L}$ themselves. Let $a, b \in \mathscr{L}$. We shall say that a and b are *simultaneously verifiable*, $a \leftrightarrow b$ in symbols, if there are elements a_1, b_1, and $c \in \mathscr{L}$ such that

$$(31)\qquad \begin{aligned} &\text{(i) } a_1, b_1, \text{ and } c \text{ are mutually orthogonal,}\\ &\text{(ii) } a = a_1 \vee c, \quad b = b_1 \vee c. \end{aligned}$$

If x and y are two observables, they are said to be *simultaneously observable* if for any two Borel subsets E and F of R^1, $x(E) \leftrightarrow y(F)$; we shall write $x \leftrightarrow y$. If A and B are two subsets of $\mathscr{L}$, we write $A \leftrightarrow B$ when $a \leftrightarrow b$ for all $a \in A$ and $b \in B$.

The analysis of these concepts depends on a study of Boolean subalgebras of $\mathscr{L}$. A subset $\mathscr{A} \subseteq \mathscr{L}$ is a *Boolean subalgebra of* $\mathscr{L}$ if it is (a) a sublogic of $\mathscr{L}$ and (b) a Boolean algebra, i.e., if

(i) $0, 1 \in \mathscr{A}$,
(ii) if $a, b \in \mathscr{A}$, then $a \vee b$, $a \wedge b$, and $a^\perp \in \mathscr{A}$,
(iii) if $a, b, c \in \mathscr{A}$, then

$$a \wedge (b \vee c) = (a \wedge b) \vee (a \wedge c),$$
$$a \vee (b \wedge c) = (a \vee b) \wedge (a \vee c).$$

$\mathscr{A}$ is said to be a *Boolean sub-σ-algebra of* $\mathscr{L}$ if it is a Boolean subalgebra, and if, for any countable sequence $a_1, a_2, \cdots \in \mathscr{A}$, $\bigvee_n a_n \in \mathscr{A}$ and $\bigwedge_n a_n \in \mathscr{A}$.

Lemma 3.7. *Let $a, b \in \mathscr{L}$, $\mathscr{L}$ being any logic. Then the following statements are equivalent:*

(32)
(a) $a \leftrightarrow b$,
(b) $a - (a \wedge b) \perp b$,
(c) $b - (a \wedge b) \perp a$,
(d) *there exists an observable x and two Borel sets A and B of the real line such that $x(A) = a$ and $x(B) = b$,*
(e) *there exists a Boolean subalgebra of $\mathscr{L}$ containing a and b.*

Proof. Suppose $a \leftrightarrow b$. We may then write $a = a_1 \vee c$, $b = b_1 \vee c$, where a_1, b_1, and c are mutually orthogonal. Clearly $c < a \wedge b$. Hence there is a d such that $c \perp d$ and $c \vee d = a \wedge b$. Since $d \perp c$ and $d < a$, $d < a - c = a_1$ (lemma 6.1). Similarly, $d < b_1$. Hence $d < a_1 \wedge b_1 = 0$. Therefore $c = a \wedge b$. Hence $a_1 = a - (a \wedge b)$. But $a_1 \perp c$ and $a_1 \perp b_1$ so that

$$a_1 = a - (a \wedge b) \perp b_1 \vee c = b.$$

This proves that (a) $\Rightarrow$ (b). By symmetry, (a) $\Rightarrow$ (c). If $a - (a \wedge b) \perp b$, then, on writing $a_1 = a - (a \wedge b)$, $b_1 = b - (a \wedge b)$ and $c = a \wedge b$, we find that $a = a_1 \vee c$ and $b = b_1 \vee c$. Since $a_1 \perp b$, we have $a_1 \perp b_1$ and $a_1 \perp c$, while, by definition, $b_1 \perp c$. Therefore $a \leftrightarrow b$, proving that (b) $\Rightarrow$ (a). (a), (b) and (c) have thus been proved equivalent. We shall complete the proof by showing (a) $\Rightarrow$ (d) $\Rightarrow$ (e) $\Rightarrow$ (a). If $a = a_1 \vee c$, $b = b_1 \vee c$, where a_1, b_1, and c are mutually orthogonal, we write $d = a_1 \vee b_1 \vee c$ and define x to be the discrete observable such that $x(\{0\}) = a_1$, $x(\{1\}) = b_1$, $x(\{2\}) = c$ and $x(\{3\}) = d^{\perp}$. Then $x(\{0,2\}) = a$, $x(\{1,2\}) = b$. This proves that (a) $\Rightarrow$ (d). Suppose that (d) is satisfied. Then $x(A - A \cap B) = a - a \wedge b$, $x(B - A \cap B) = b - a \wedge b$. Writing $a_1 = a - a \wedge b$, $a_2 = a \wedge b$, $a_3 = b - a \wedge b$, $a_4 = (a \vee b)^{\perp}$, we see that the a_i are mutually orthogonal and $a_1 \vee a_2 \vee a_3 \vee a_4 = 1$. If $\mathscr{A}$ is the collection of elements which are the lattice sums

$$a_{i_1} \vee a_{i_2} \vee \cdots \vee a_{i_k} \quad (k \leq 4, 1 \leq i_1 \leq i_2 \leq \cdots \leq i_k \leq 4),$$

then it is easy to verify that $\mathscr{A}$ is a Boolean subalgebra of $\mathscr{L}$. Since $a, b \in \mathscr{A}$, we see that (d) $\Rightarrow$ (e). Finally, let (e) be satisfied, and let $\mathscr{A}$ be a Boolean subalgebra of $\mathscr{L}$ containing a and b. Clearly $[a - (a \wedge b)] \wedge b = 0$. But, as $a - (a \wedge b)$, b and $b^{\perp}$ are in the Boolean algebra $\mathscr{A}$,

$$\begin{aligned} a - (a \wedge b) &= \{(a - (a \wedge b)) \wedge b\} \vee \{(a - (a \wedge b)) \wedge b^{\perp}\} \\ &= \{a - (a \wedge b)\} \wedge b^{\perp} \\ &< b^{\perp}, \end{aligned}$$

showing that $a - (a \wedge b) \perp b$. This shows that (e) $\Rightarrow$ (b) and hence that (e) $\Rightarrow$ (a).

Corollary 3.8. *If* $a \leftrightarrow b$, *the elements* a_1, b_1, and c *of* (31) *are uniquely determined. In fact,* $c = a \wedge b$, $a_1 = a - a \wedge b$, *and* $b_1 = b - a \wedge b$.

Lemma 3.7 shows that in the presence of simultaneous verifiability of two elements of $\mathscr{L}$, one can operate with the statements corresponding to these elements as if they were classical. We shall now take up the question of generalizing this fact to observables. Our aim is to prove that *simultaneous observability* is essentially a characteristic of classical systems. We shall deduce, as a corollary, that, if $x, y, z, \cdots$ are observables, associated with a logic, they are mutually simultaneously observable if and only if they are all functions of a single observable. This was first proved by von Neumann (cf. [1]) when the logic in question was a standard one. For a general logic, it was proved by Varadarajan [1].

Theorem 3.9. † *Let* $\mathscr{L}$ *be any logic and* $\{x_\lambda\}_{\lambda \in D}$ *a family of observables. Suppose that* $x_\lambda \leftrightarrow x_{\lambda'}$ *for all* $\lambda, \lambda' \in D$. *Then, there exists a space* X, *a*

---

† The definition of a logic given in Varadarajan [1] does not require it to be a lattice. It was pointed out to us by Dr. S. P. Gudder [1] that theorem 3.9 does not remain valid for these more general classes of partially ordered sets considered in our paper [1].

σ-algebra $\mathscr{B}$ of subsets of X, real valued $\mathscr{B}$-measurable functions g_λ on $X(\lambda \in D)$, and a σ-homomorphism τ of $\mathscr{B}$ into $\mathscr{L}$ such that

$$\tau(g_\lambda{}^{-1}(E)) = x_\lambda(E) \tag{33}$$

for all $\lambda \in D$ and $E \in \mathscr{B}(R^1)$. Suppose further that either $\mathscr{L}$ is separable in the sense that every Boolean sub-σ-algebra of $\mathscr{L}$ is separable, or that D is countable. Then there exists an observable x and real valued Borel functions f_λ of a real variable such that for all $\lambda \in D$,

$$x_\lambda = f_\lambda \circ x. \tag{34}$$

The proof of this theorem (cf. Varadarajan [1]) depends on a series of lemmas. $\mathscr{L}$ is a fixed logic in these lemmas.

Lemma 3.10. *Let $b, a_1, a_2, \cdots$ be elements of $\mathscr{L}$. If $b \leftrightarrow a_n$ for all n, then*

$$\begin{aligned} &b \leftrightarrow a_n{}^\perp, \\ &b \leftrightarrow \bigvee_n a_n, \\ &b \leftrightarrow \bigwedge_n a_n. \end{aligned} \tag{35}$$

Moreover, we then have the distributivity laws

$$\begin{aligned} b \wedge \Big(\bigvee_n a_n\Big) &= \bigvee_n (b \wedge a_n), \\ b \vee \Big(\bigwedge_n a_n\Big) &= \bigwedge_n (b \vee a_n). \end{aligned} \tag{36}$$

Proof. We use the criteria of (32). It follows from the equivalence of (a) and (d) of (32) that if $c \leftrightarrow d$, then any two of the four elements c, $c^\perp$, d, $d^\perp$ are simultaneously verifiable. This proves the first of the three relations of (35). Now, we come to the proof of the rest of (35). Write $a = \bigvee_m a_m$. Since $b \wedge a_m < b \wedge a$, $b-(b \wedge a) < b-(b \wedge a_m)$ for all m. As $b \leftrightarrow a_m$ for all m, it follows from (c) of (32) that

$$b-(b \wedge a) < b-(b \wedge a_m) < a_m{}^\perp.$$

Hence $b-(b \wedge a) < \bigwedge_m a_m{}^\perp = a^\perp$ which proves that $b \leftrightarrow a$. Changing all the a_n into $a_n{}^\perp$ we deduce from the relation $b \leftrightarrow \bigvee_n a_n$ that $b \leftrightarrow \bigwedge_n a_n$. Next we come to (36). Once again it is enough to prove only the first of the two relations. Clearly $\bigvee_n (b \wedge a_n) < b \wedge a$, where $a = \bigvee_n a_n$. Let $d < b \wedge a$ be such that $d \perp \bigvee_n (b \wedge a_n)$ and $(\bigvee_n (b \wedge a_n)) \vee d = b \wedge a$. Since $d \perp b \wedge a_n$ and $d < b$, we deduce from the relation $b \leftrightarrow a_n$ that $d < b-(b \wedge a_n)$ (by lemma 6.1) and hence, as $b \leftrightarrow a_n$, we conclude that $d \perp a_n$, from (c) of (32). Since this is true for all n, $d \perp a$. As $d < a$, this means $d = 0$, proving the first equation of (36).

Consider the collection $\mathfrak{M}$ of sublogics of $\mathscr{L}$, partially ordered by inclusion. If $\{\mathscr{L}_\alpha\}_{\alpha \in A}$ is any family of sublogics of $\mathscr{L}$, it is obvious that

$\bigcap_\alpha \mathscr{L}_\alpha$ is also a sublogic of $\mathscr{L}$. From this it follows that if $\{\mathscr{L}_\alpha\}_{\alpha\in A}$ is any family of sublogics of $\mathscr{L}$, there exists a smallest sublogic containing all the $\mathscr{L}_\alpha$. We denote it by $\bigvee_\alpha \mathscr{L}_\alpha$. $\mathfrak{M}$ is thus a *complete* lattice. If each $\mathscr{L}_\alpha$ is a Boolean subalgebra of $\mathscr{L}$, it is however not true in general that $\bigvee_\alpha \mathscr{L}_\alpha$ is a Boolean subalgebra. We shall now prove that this is so if and only if the relations $\mathscr{L}_\alpha \leftrightarrow \mathscr{L}_{\alpha'}$ are satisfied for all $\alpha, \alpha' \in A$. We begin by considering the case when there are only two Boolean subalgebras involved.

Lemma 3.11. *Let $\mathscr{R}_1$ and $\mathscr{R}_2$ be two Boolean subalgebras of $\mathscr{L}$. Then $\mathscr{R}_1 \leftrightarrow \mathscr{R}_2$ if and only if $\mathscr{R}_1 \vee \mathscr{R}_2$ is a Boolean subalgebra of $\mathscr{L}$.*

Proof. Suppose that $\mathscr{R}_1 \vee \mathscr{R}_2$ is a Boolean subalgebra of $\mathscr{L}$. Then $\mathscr{R}_1 \leftrightarrow \mathscr{R}_2$ in view of (e) of (32). We shall now prove the converse. We shall first consider the case when the Boolean subalgebras are finite. Let $\mathscr{F}_i$ $(i=1, 2)$ be *finite* Boolean subalgebras of $\mathscr{L}$ such that $\mathscr{F}_1 \leftrightarrow \mathscr{F}_2$. Then, their Stone spaces (cf. Chapter I) are finite and hence there are elements $a_1, \cdots, a_r \in \mathscr{F}_1$ and $b_1, \cdots, b_s \in \mathscr{F}_2$ such that (i) $a_i \perp a_{i'}$, $i \neq i'$, $b_j \perp b_{j'}$, $j \neq j'$; (ii) $a_1 \vee a_2 \vee \cdots \vee a_r = b_1 \vee b_2 \vee \cdots \vee b_s = 1$; (iii) $\mathscr{F}_1$ is the collection of all lattice sums $a_{i_1} \vee a_{i_2} \vee \cdots \vee a_{i_p}$ $(1 \leq i_1 \leq i_2 \leq \cdots \leq i_p \leq r)$ and $\mathscr{F}_2$ is the collection of all lattice sums $b_{j_1} \vee b_{j_2} \vee \cdots \vee b_{j_q}$ $(1 \leq j_1 \leq j_2 \leq \cdots \leq j_q \leq s)$. Write

$$c_{ij} = a_i \wedge b_j \qquad (i = 1, 2, \cdots, r, \quad j = 1, 2, \cdots, s). \tag{37}$$

Obviously $c_{ij} \perp c_{i'j'}$ unless $i=i'$ and $j=j'$. Since $a_i \leftrightarrow b_j$ for all i and j, we infer from (36) that

$$\begin{aligned} \bigvee_{j=1}^{s} c_{ij} &= a_i, \\ \bigvee_{i=1}^{r} c_{ij} &= b_j. \end{aligned} \tag{38}$$

Consequently,

$$\bigvee_{i=1}^{r} \bigvee_{j=1}^{s} c_{ij} = 1. \tag{39}$$

If $\mathscr{F}$ denotes the collection of all lattice sums of the c_{ij}, then $\mathscr{F}$ is a Boolean subalgebra of $\mathscr{L}$ and (38) implies that $\mathscr{F}_i \subseteq \mathscr{F}$ for $i=1, 2$. The formulas (37) show that if $\mathscr{F}'$ is any sublogic of $\mathscr{L}$ containing both $\mathscr{F}_1$ and $\mathscr{F}_2$, then $\mathscr{F} \subseteq \mathscr{F}'$. Clearly, therefore, $\mathscr{F} = \mathscr{F}_1 \vee \mathscr{F}_2$. Note that $\mathscr{F}_1 \vee \mathscr{F}_2$ is also finite.

Let $\mathscr{R}_i$ $(i=1, 2)$ be Boolean subalgebras of $\mathscr{L}$ such that $\mathscr{R}_1 \leftrightarrow \mathscr{R}_2$. Let $\mathscr{R}$ be defined by

$$\mathscr{R} = \bigcup (\mathscr{F}_1 \vee \mathscr{F}_2), \tag{40}$$

where the union is over all pairs $(\mathscr{F}_1, \mathscr{F}_2)$ of finite Boolean subalgebras of $\mathscr{L}$ such that $\mathscr{F}_i \subseteq \mathscr{R}_i$, $i=1, 2$. We assert that $\mathscr{R}$ is a Boolean subalgebra of $\mathscr{L}$ containing $\mathscr{R}_1$ and $\mathscr{R}_2$. If $a \in \mathscr{R}$, then it is obvious that $a^\perp \in \mathscr{R}$. Clearly, $0, 1 \in \mathscr{R}$. Suppose now that $a, b \in \mathscr{R}$. Then there are finite Boolean subalgebras $\mathscr{F}_1, \mathscr{F}_1', \mathscr{F}_2, \mathscr{F}_2'$ of $\mathscr{L}$ such that $\mathscr{F}_1, \mathscr{F}_1'$ are $\subseteq \mathscr{R}_1$, $\mathscr{F}_2, \mathscr{F}_2'$ are $\subseteq \mathscr{R}_2$, $a \in \mathscr{F}_1 \vee \mathscr{F}_2$, and $b \in \mathscr{F}_1' \vee \mathscr{F}_2'$. Since $\mathscr{F}_1$ and $\mathscr{F}_1'$ are contained in the Boolean algebra $\mathscr{R}_1$, $\mathscr{F}_1 \leftrightarrow \mathscr{F}_1'$ and hence $\mathscr{F}_1 \vee \mathscr{F}_1' = \mathscr{F}_1''$ is finite and contained in $\mathscr{R}_1$. Similarly, $\mathscr{F}_2 \vee \mathscr{F}_2' = \mathscr{F}_2''$ is finite and contained in $\mathscr{R}_2$. As $\mathscr{R}_1 \leftrightarrow \mathscr{R}_2$, $\mathscr{F}_1'' \leftrightarrow \mathscr{F}_2''$. Therefore $\mathscr{F}_1'' \vee \mathscr{F}_2''$ is a finite Boolean subalgebra of $\mathscr{L}$ contained in $\mathscr{R}$. Let us write $\mathscr{F} = \mathscr{F}_1'' \vee \mathscr{F}_2''$. As $\mathscr{F}_1, \mathscr{F}_1', \mathscr{F}_2$, and $\mathscr{F}_2'$ are contained in $\mathscr{F}$, it follows that $\mathscr{F}_1 \vee \mathscr{F}_2 \subseteq \mathscr{F}$ and $\mathscr{F}_1' \vee \mathscr{F}_2' \subseteq \mathscr{F}$. Thus, $a, b \in \mathscr{F}$. Therefore $a \vee b$ and $a \wedge b$ are in $\mathscr{F}$. This proves that $a \vee b$ and $a \wedge b$ are in $\mathscr{R}$.

We shall next establish the distributivity laws. Let a_1, a_2, a_3 be three elements of $\mathscr{R}$ and (for $j = 1, 2, 3$) let $a_j \in \mathscr{F}_j^{(1)} \vee \mathscr{F}_j^{(2)}$, where $\mathscr{F}_j^{(k)}$ is a finite Boolean subalgebra of $\mathscr{L}$ contained in $\mathscr{R}_k$ ($j = 1, 2, 3$, $k = 1, 2$). Since $\mathscr{F}_j^{(k)} \subseteq \mathscr{R}_k$ for all j, $\bigvee_{j=1}^3 \mathscr{F}_j^{(k)}$ is a finite Boolean subalgebra of $\mathscr{R}_k$. Let it be denoted by $\mathscr{F}^{(k)}$. Obviously $\mathscr{F}^{(1)} \leftrightarrow \mathscr{F}^{(2)}$, and hence $\mathscr{F} = \mathscr{F}^{(1)} \vee \mathscr{F}^{(2)}$ is a finite Boolean subalgebra of $\mathscr{L}$. Since $\mathscr{F}_j^{(k)} \subseteq \mathscr{F}$ for all j and k, $\mathscr{F}_j^{(1)} \vee \mathscr{F}_j^{(2)} \subseteq \mathscr{F}$ for all j so that $a_j \in \mathscr{F}$ for all j. As $\mathscr{F}$ is a Boolean subalgebra of $\mathscr{L}$, we have:

$$a_1 \wedge (a_2 \vee a_3) = (a_1 \wedge a_2) \vee (a_1 \wedge a_3),$$
$$a_1 \vee (a_2 \wedge a_3) = (a_1 \vee a_2) \wedge (a_1 \vee a_3).$$

This proves that $\mathscr{R}$ is a Boolean subalgebra of $\mathscr{L}$.

We observe finally that $\mathscr{R}_1$ and $\mathscr{R}_2$ are contained in $\mathscr{R}$. If $a \in \mathscr{R}_1$, we have $a \in \mathscr{F}_1 \vee \mathscr{F}_2$, where $\mathscr{F}_1 = \{0, a, a^\perp, 1\}$ and $\mathscr{F}_2 = \{0, 1\}$. Hence $\mathscr{R}_1 \subseteq \mathscr{R}$. Similarly, $\mathscr{R}_2 \subseteq \mathscr{R}$. $\mathscr{R}_1 \vee \mathscr{R}_2$ is thus a Boolean subalgebra of $\mathscr{L}$. The lemma is completely proved.

Corollary 3.12. *If $\mathscr{R}$ is defined by* (40), *then $\mathscr{R} = \mathscr{R}_1 \vee \mathscr{R}_2$. If $\mathscr{C}$ is any subset of $\mathscr{L}$ such that $\mathscr{C} \leftrightarrow \mathscr{R}_k$, $k = 1, 2$, then $\mathscr{C} \leftrightarrow \mathscr{R}_1 \vee \mathscr{R}_2$.*

Proof. If $\mathscr{F}_k$ is a finite Boolean subalgebra of $\mathscr{R}_k$ ($k = 1, 2$) and if $\mathscr{R}'$ is a sublogic of $\mathscr{L}$ which contains both $\mathscr{R}_1$ and $\mathscr{R}_2$, $\mathscr{F}_1 \vee \mathscr{F}_2 \subseteq \mathscr{R}'$. Hence $\mathscr{R} \subseteq \mathscr{R}'$. This shows that $\mathscr{R}_1 \vee \mathscr{R}_2$ coincides with $\mathscr{R}$. To prove the second assertion, let $c \in \mathscr{C}$. Let $\mathscr{F}_k \subseteq \mathscr{R}_k$ be finite Boolean subalgebras of $\mathscr{L}$. From (37) and (35) we may conclude that $c \leftrightarrow \mathscr{F}_1 \vee \mathscr{F}_2$. Equation (40) shows then that $c \leftrightarrow \mathscr{R}$. This proves that $\mathscr{C} \leftrightarrow \mathscr{R}_1 \vee \mathscr{R}_2$.

Corollary 3.13. *Let $\mathscr{R}_i$ ($i = 1, 2, \cdots, s$) be Boolean subalgebras of $\mathscr{L}$ such that $\mathscr{R}_i \leftrightarrow \mathscr{R}_j$ for all $i, j = 1, 2, \cdots, s$. Then $\bigvee_{i=1}^s \mathscr{R}_i$ is a Boolean subalgebra of $\mathscr{L}$. If $\mathscr{C}$ is any subset of $\mathscr{L}$ such that $\mathscr{C} \leftrightarrow \mathscr{R}_i$ for all $i = 1, 2, \cdots, s$, then $\mathscr{C} \leftrightarrow \bigvee_{i=1}^s \mathscr{R}_i$.*

Proof. The assertions are proved by induction on s. For $s = 2$, they follow from lemma 3.11 and corollary 3.12. Assume now that they are valid for all $s \leq s_0$ and let $s = s_0 + 1$. Since $\mathscr{R}_i \leftrightarrow \mathscr{R}_j$ for all $i, j \leq s_0 + 1$, $\mathscr{R} = \bigvee_{i=1}^{s_0} \mathscr{R}_i$ is a Boolean subalgebra of $\mathscr{L}$ and $\mathscr{R}_{s_0+1} \leftrightarrow \mathscr{R}$ by the induction hypothesis. Hence, $\mathscr{R}_{s_0+1} \vee \mathscr{R}$ is a Boolean subalgebra of $\mathscr{L}$. Suppose now that c is an element of $\mathscr{L}$ such that $c \leftrightarrow \mathscr{R}_i$, $i \leq s_0 + 1$. Then, by the induction hypothesis, $c \leftrightarrow \mathscr{R}$. Hence we conclude from corollary 3.12 that $c \leftrightarrow \bigvee_{i=1}^{s_0+1} \mathscr{R}_i$. This completes the induction and the proof of the corollary.

Lemma 3.14. *Let $\{\mathscr{R}_\lambda\}_{\lambda \in D}$ be a family of Boolean subalgebras of $\mathscr{L}$ such that $\mathscr{R}_\lambda \leftrightarrow \mathscr{R}_{\lambda'}$, for all $\lambda, \lambda' \in D$. Then there exists a Boolean subalgebra $\mathscr{R}$ of $\mathscr{L}$ such that $\mathscr{R}_\lambda \subseteq \mathscr{R}$ for all $\lambda \in D$.*

Proof. Let F be a finite subset of D. By corollary 3.13, $\bigvee_{\lambda \in F} \mathscr{R}_\lambda$ is a Boolean subalgebra of $\mathscr{L}$. Let us write

$$\mathscr{R}(F) = \bigvee_{\lambda \in F} \mathscr{R}_\lambda. \tag{41}$$

Obviously, $\mathscr{R}(F) \subseteq \mathscr{R}(F')$ if $F \subseteq F'$. Define $\mathscr{R}$ by

$$\mathscr{R} = \bigcup_{\substack{F \subseteq D \\ F \text{ finite}}} \mathscr{R}(F). \tag{42}$$

Obviously, $\mathscr{R}_\lambda \subseteq \mathscr{R}$ for all $\lambda \in D$. We shall now complete the proof by showing that $\mathscr{R}$ is a Boolean subalgebra of $\mathscr{L}$. It is enough to prove that any finite subset of $\mathscr{R}$ is contained in a Boolean subalgebra of $\mathscr{L}$ contained in $\mathscr{R}$. If $a_1, a_2, \cdots, a_r$ are elements of $\mathscr{R}$ there are finite subsets F_1, $F_2, \cdots, F_r$ of D such that $a_i \in \mathscr{R}(F_i)$. If $F = \bigcup_i F_i$, then $a_i \in \mathscr{R}(F)$ for all i. This proves that $\mathscr{R}$ is a Boolean subalgebra of $\mathscr{L}$.

Corollary 3.15. *Let $\{\mathscr{R}_\lambda\}_{\lambda \in D}$ be a family of Boolean sub-σ-algebras of $\mathscr{L}$ such that $\mathscr{R}_\lambda \leftrightarrow \mathscr{R}_\lambda'$ for all $\lambda, \lambda' \in D$. Then, there exists a Boolean sub-σ-algebra $\mathscr{R}$ of $\mathscr{L}$ such that $\mathscr{R}_\lambda \subseteq \mathscr{R}$ for all $\lambda \in D$.*

Proof. By lemma 3.14 there are Boolean subalgebras $\mathscr{R}'$ such that $\mathscr{R}_\lambda \subseteq \mathscr{R}'$ for all λ. By Zorn's lemma we can choose a maximal such, say $\mathscr{R}$. We claim that $\mathscr{R}$ is a Boolean σ-algebra. It is enough to prove that if $a_1, a_2, \cdots$ is a sequence of elements of $\mathscr{R}$, then $a = \bigvee_m a_m$ is also in $\mathscr{R}$. Since $a_m \in \mathscr{R}$, $b \leftrightarrow a_m$ for all $b \in \mathscr{R}$. Hence $a \leftrightarrow \mathscr{R}$ by (35). If $\mathscr{A} = \{0, a, a^\perp, 1\}$ then $\mathscr{A}$ is a Boolean subalgebra of $\mathscr{L}$ and $\mathscr{A} \leftrightarrow \mathscr{R}$. Hence $\mathscr{A} \vee \mathscr{R}$ is a Boolean subalgebra of $\mathscr{L}$ containing $\mathscr{R}$. Since $\mathscr{R}$ is maximal, $\mathscr{A} \vee \mathscr{R} = \mathscr{R}$. Therefore $a \in \mathscr{R}$.

For any observable x associated with $\mathscr{L}$ we write

$$\mathscr{R}(x) = \{x(E) : E \in \mathscr{B}(R^1)\}. \tag{43}$$

We shall call $\mathscr{R}(x)$ the *range* of x.

Lemma 3.16. *Let $\mathscr{R}$ be a Boolean sub-σ-algebra of $\mathscr{L}$. Then, in order that $\mathscr{R} = \mathscr{R}(x)$ for some observable x, it is necessary and sufficient that $\mathscr{R}$ be separable.*

Proof. If $\mathscr{R}=\mathscr{R}(x)$ and $\mathscr{D}=\{x(E) : E$ an open interval with rational end points$\}$, then $\mathscr{D}$ is a countable subset of $\mathscr{R}$ and $\mathscr{R}$ is the smallest sub-σ-algebra of itself containing $\mathscr{D}$. $\mathscr{R}$ is thus separable. Conversely, let $\mathscr{R}$ be a separable Boolean sub-σ-algebra of $\mathscr{L}$. By theorem 1.3, there exists a space X, a σ-algebra $\mathscr{B}$ of subsets of X, and a σ-homomorphism h of $\mathscr{B}$ onto $\mathscr{R}$. Let $\{A_1, A_2, \cdots\}$ be a sequence of subsets of X such that $A_n \in \mathscr{B}$ for all n and $\{h(A_n) : n=1, 2, \cdots\}$ generates $\mathscr{R}$. Let $\mathscr{B}_0$ be the sub-σ-algebra of $\mathscr{B}$ generated by $\{A_n : n=1, 2, \cdots\}$. By theorem 1.6(i), there exists a σ-homomorphism u of $\mathscr{B}(R^1)$ onto $\mathscr{B}_0$. Define x by

$$x(E) = h(u(E)) \qquad (E \in \mathscr{B}(R^1)). \tag{44}$$

It is obvious that x is an observable and $\mathscr{R}=\mathscr{R}(x)$.

Now to the proof of theorem 3.9. Suppose that x is an observable and $x_\lambda = f_\lambda \circ x$ $(\lambda \in D)$, where each f_λ is a real Borel function on R^1. Then $\mathscr{R}(x_\lambda) \subseteq \mathscr{R}(x)$ and hence $\mathscr{R}(x_\lambda) \leftrightarrow \mathscr{R}(x_{\lambda'})$ for all $\lambda, \lambda' \in D$. This proves that $x_\lambda \leftrightarrow x_{\lambda'}$, for all $\lambda, \lambda' \in D$. Conversely, suppose that $\{x_\lambda\}$ is a family of observables and that $x_\lambda \leftrightarrow x_{\lambda'}$ for all $\lambda, \lambda' \in D$. By corollary 3.15, there exists a Boolean sub-σ-algebra $\mathscr{R}'$ of $\mathscr{L}$ such that $\mathscr{R}(x_\lambda) \subseteq \mathscr{R}'$ for all λ. By theorem 1.3, there exists a space X, a σ-algebra $\mathscr{B}$ of subsets of X, and a σ-homomorphism τ of $\mathscr{B}$ onto $\mathscr{R}'$. It is now clear that for each $\lambda \in D$, there exists, by theorem 1.4, a real valued $\mathscr{B}$-measurable function g_λ on X such that, for all $E \in \mathscr{B}(R^1)$,

$$x_\lambda(E) = \tau[g_\lambda{}^{-1}(E)] \qquad (\lambda \in D).$$

This proves the equation (33). For proving (34) let $\mathscr{R}$ be the smallest sub-σ-algebra of $\mathscr{R}'$ containing all the $\mathscr{R}(x_\lambda)$. If $\mathscr{L}$ is separable, so is $\mathscr{R}$. On the other hand, if D is countable, and $\mathscr{D}_\lambda$ is a countable set generating $\mathscr{R}(x_\lambda)$, $\bigcup_\lambda \mathscr{D}_\lambda$ is a countable set generating $\mathscr{R}$. Hence $\mathscr{R}$ is separable in either case. By lemma 3.16 there exists an observable x such that $\mathscr{R}(x)=\mathscr{R}$. As $\mathscr{R}(x_\lambda) \subseteq \mathscr{R}(x)$, we conclude from theorem 1.6(ii) that there exists a real Borel function f_λ of a real variable such that $x_\lambda = f_\lambda \circ x$. This proves (34) and completes the proof of the entire theorem.

Remark. It is obvious that if the x_λ satisfy either (33) or (34), $x_\lambda \leftrightarrow x_{\lambda'}$ for all $\lambda, \lambda' \in D$. Theorem 3.9 therefore tells us that the classical logics are precisely those for which any two observables are simultaneously observable. As soon as $\mathscr{L}$ is not a Boolean algebra, it follows that the properties of a physical system which are embodied in the logic $\mathscr{L}$ have too complex a structure to be exhausted by a single set of simultaneously observable quantities. In general, when certain propositions are experimentally verified, certain others cannot be so at the same time.

6. FUNCTIONS OF SEVERAL OBSERVABLES

The characterization given in theorem 3.9 of simultaneous observability enables us to construct a calculus of functions of several observables which are mutually simultaneously observable.

Theorem 3.17. *Let* $\mathscr{L}$ *be any logic and* $x_1, x_2, \cdots, x_n$ *observables such that* $x_i \leftrightarrow x_j$ *for all* i, j. *Then there exists one and only one* σ*-homomorphism* τ *of* $\mathscr{B}(R^n)$ *into* $\mathscr{L}$ *such that for all* $E \in \mathscr{B}(R^1)$ *and all* $i = 1, 2, \cdots, n$,

$$x_i(E) = \tau(\pi_i^{-1}(E)), \tag{45}$$

where π_i *is the projection* $(t_1, t_2, \cdots, t_n) \rightarrow t_i$ *of* R^n *on* R^1. *If* g *is any real valued Borel function on* R^n,

$$g \circ (x_1, x_2, \cdots, x_n) : E \rightarrow \tau(g^{-1}(E)) \tag{46}$$

is an observable. If $g_1, g_2, \cdots, g_k$ *are real valued Borel functions on* R^n *and* $y_i = g_i \circ (x_1, \cdots, x_n)$, *then* $y_1, \cdots, y_k$ *are simultaneously observable, and for any real valued Borel function* h *on* R^k,

$$h \circ (y_1, \cdots, y_k) = (h(g_1, \cdots, g_k)) \circ (x_1, \cdots, x_n), \tag{47}$$

where $h(g_1, \cdots, g_k)$ *is the function*

$$\mathbf{t} = (t_1, \cdots, t_n) \rightarrow h(g_1(\mathbf{t}), \cdots, g_k(\mathbf{t})).$$

Proof. By theorem 3.9, there exists an observable x and real valued Borel functions $f_1, \cdots, f_n$ of a real variable such that $x_i = f_i \circ x$ for all i. Let the Borel map $\mathbf{f}$ of R^1 into R^n be defined by

$$\mathbf{f} : t \rightarrow (f_1(t), \cdots, f_n(t)).$$

If we define, for $M \in \mathscr{B}(R^n)$,

$$\tau(M) = x(\mathbf{f}^{-1}(M)), \tag{48}$$

then τ is a σ-homomorphism of $\mathscr{B}(R^n)$ into $\mathscr{L}$ which satisfies (45). If τ and τ' are two σ-homomorphisms of $\mathscr{B}(R^n)$ into $\mathscr{L}$ satisfying (45), then

$$\tau(M) = \tau'(M)$$

whenever M is of the form $E_1 \times E_2 \times \cdots \times E_n$, the E_j being Borel subsets of R^1. Since the class of such sets generates $\mathscr{B}(R^n)$, it follows that $\tau = \tau'$. The fact that $g \circ (x_1, \cdots, x_n)$ is an observable is trivial. Note that its range is contained in the range $\mathscr{R}$ of τ which is a Boolean sub-σ-algebra of $\mathscr{L}$.

Suppose now that $y_i = g_i \circ (x_1, \cdots, x_n)$. Then the range of y_i is contained in $\mathscr{R}$ and hence $y_1, \cdots, y_k$ are simultaneously observable. Let $\mathbf{g}$ be the map $\mathbf{t} \to (g_1(\mathbf{t}), \cdots, g_k(\mathbf{t}))$ of R^n into R^k. Then

$$\gamma : M \to \tau(\mathbf{g}^{-1}(M))$$

is a σ-homomorphism of $\mathscr{B}(R^k)$ into $\mathscr{L}$ such that

$$\gamma(\pi_i{}^{-1}(E)) = \tau(g_i{}^{-1}(E)) = y_i(E),$$

π_i being the map $(s_1, s_2, \cdots, s_k) \to s_i$ of R^k on R^1 $(i = 1, 2, \cdots, k)$. Therefore we conclude from the uniqueness of the map τ in (45) that, for

$$u = h \circ (y_1, y_2, \cdots, y_k),$$

$$u(E) = \gamma(h^{-1}(E)) \qquad (E \in \mathscr{B}(R^1)).$$

But $\gamma(h^{-1}(E)) = \tau(\mathbf{g}^{-1}(h^{-1}(E)))$, so that we may conclude that

$$h \circ (y_1, \cdots, y_k) = (h(g_1, \cdots, g_k)) \circ (x_1, \cdots, x_n).$$

This proves the theorem.

Theorem 3.17 enables us to define the *joint distributions* of $x_1, \ldots, x_n$. Let $p(a \to p(a))$ be a state of $\mathscr{L}$ and $x_1, \cdots, x_n$ observables such that $x_i \leftrightarrow x_j$ for all i, j. We may then clearly define the probability measure $P^p_{x_1, \cdots, x_n}$ on $\mathscr{B}(R^n)$ by

$$P^p_{x_1, \cdots, x_n}(M) = p(\tau(M)). \tag{49}$$

$P^p_{x_1, \cdots, x_n}$ is called the joint *probability distribution* of $(x_1, \cdots, x_n)$ *in the state p*. It follows easily from (47) that for Borel $M \subseteq R^k$,

$$P^p_{y_1, \cdots, {}_1 y_k}(M) = P^p_{x_1, \cdots, x_n}(\mathbf{g}^{-1}(M)), \tag{50}$$

where $\mathbf{g}$ is the map

$$\mathbf{t} = (t_1, \cdots, t_n) \to (g_1(\mathbf{t}), \cdots, g_k(\mathbf{t})).$$

We see from (50) that the rules for the calculation of the joint probability distributions are the standard ones of probability theory.

7. THE CENTER OF A LOGIC

We begin by recalling the discussion of Chapter II in which we singled out the geometries among the modular lattices of finite rank by demanding that their centers be trivial. We shall now do the same for logics. The irreducible logics would then have properties very far removed from those of Boolean algebras and would consequently serve as plausible models for the logics of atomic systems.

Let $\mathscr{L}$ be any logic. We define the *center* of $\mathscr{L}$ to be the set $\mathscr{C}$, where

$$(51) \qquad \mathscr{C} = \{a : a \in \mathscr{L}, \quad a \leftrightarrow b, \quad \text{for all } b \in \mathscr{L}\}.$$

$\mathscr{C}$ is nonempty since 0 and 1 are in $\mathscr{C}$. The elements of $\mathscr{C}$ are thus simultaneously verifiable with every element of $\mathscr{L}$. If x is any observable with $\mathscr{R}(x) \subseteq \mathscr{C}$, then for any observable y, $x \leftrightarrow y$. Such observables x are called *central*. $\mathscr{L}$ is said to be *irreducible* if $\mathscr{C} = \{0,1\}$. Clearly, $\mathscr{L}$ is irreducible if and only if the only central observables are the constants.

Theorem 3.18. *Let $\mathscr{L}$ be any logic and $\mathscr{C}$ its center. Then $\mathscr{C}$ is a Boolean sub-σ-algebra of $\mathscr{L}$.*

Proof. The equations (35) and (36) reveal immediately that $\mathscr{C}$ is a Boolean subalgebra of $\mathscr{L}$. If $a_1, \cdots, a_r, \cdots$ is a sequence of elements in $\mathscr{C}$, then $b \leftrightarrow a_r$ for all $r (b \in \mathscr{L})$ and so $\bigvee_n a_n \leftrightarrow b$ by (35). Hence $\bigvee_n a_n \in \mathscr{C}$ also. This proves that $\mathscr{C}$ is a Boolean sub-σ-algebra of $\mathscr{L}$.

Since the Boolean σ-algebras are relatively familiar, one may ask whether there is a decomposition of $\mathscr{L}$ over $\mathscr{C}$ analogous to that for modular lattices of finite rank. There is no difficulty in doing this when $\mathscr{C}$ is *discrete*, i.e., when there exists an at most denumerable set $\{a_n\}_{n \in D}$ of mutually orthogonal elements of $\mathscr{C}$ such that (i) $\bigvee_n a_n = 1$; (ii) $\mathscr{C}$ consists precisely of all the lattice sums $\bigvee_{n \in Z} a_n$, where Z is an arbitrary subset of D. The a_n are called the *atoms* of $\mathscr{C}$. If $\mathscr{C}$ is not discrete, then such decompositions seem much more difficult to obtain.

On the other hand, let us define $\mathscr{C}$ to be *continuous* if, given any $c \in \mathscr{C}$ which is nonzero, there exists a nonzero $c_1 \in \mathscr{C}$ such that $c_1 < c$, $c_1 \neq c$. We shall prove that when $\mathscr{C}$ is continuous and separable, $\mathscr{L}$ does not have any pure states.

Theorem 3.19. *Let $\mathscr{L}$ be any logic and let $\mathscr{C}$ be its center. Let $\mathscr{P}$ be the set of pure states of $\mathscr{L}$. Suppose that $\mathscr{C}$ is discrete and that $\{a_n : n \in D\}$ is the family of atoms of $\mathscr{C}$. Then for each $n \in D$, $\mathscr{L}[0,a_n] = \mathscr{L}_n$ is an irreducible logic. If p is a pure state of $\mathscr{L}_n$ and we define $p^\sim$ on $\mathscr{L}$ by*

$$(52) \qquad p^\sim(a) = p(a \wedge a_n) \qquad (a \in \mathscr{L}),$$

then $p^\sim$ is a pure state of $\mathscr{L}$. If

$$(53) \qquad \mathscr{P}_n = \{p^\sim : p \text{ a pure state of } \mathscr{L}_n\},$$

then the $\{\mathscr{P}_n : n \in D\}$ are disjoint subsets of $\mathscr{P}$ and

$$(54) \qquad \mathscr{P} = \bigcup_{n \in D} \mathscr{P}_n.$$

If $\mathscr{C}$ is continuous and separable, $\mathscr{P}$ is empty, i.e., $\mathscr{L}$ has no pure states.

Proof. We have already noted that for any n, $\mathscr{L}[0,a_n]=\mathscr{L}_n$ is a logic (corollary 3.3). We prove that $\mathscr{L}_n$ is irreducible. In fact, let $c<a_n$ be such that $c \leftrightarrow \mathscr{L}_n$. Clearly, $c \leftrightarrow a$ for all $a<a_n$. On the other hand, as $a_n \in \mathscr{C}$, $a_n \leftrightarrow b$ for any $b \in \mathscr{L}$, so that $b=(b \wedge a_n) \vee (b \wedge a_n{}^{\perp})$ (cf. (36)). Now $c \leftrightarrow b \wedge a_n$ as $b \wedge a_n < a_n$ while $c \leftrightarrow b \wedge a_n{}^{\perp}$ since $c \perp b \wedge a_n{}^{\perp}$. Therefore (35) enables us to conclude that $c \leftrightarrow b$. This shows that $c \in \mathscr{C}$ and hence that $c=0$ or $c=a_n$. This proves the irreducibility of $\mathscr{L}_n$.

Suppose next that p is a pure state of $\mathscr{L}_n$. Equation (52) defines $p^{\sim}$ as a function on $\mathscr{L}$. From the equations (35) and (36) it follows that $p^{\sim}$ is a state of $\mathscr{L}$. We claim that $p^{\sim}$ is a pure state of $\mathscr{L}$. Suppose that $p^{\sim}=cp_1+(1-c)p_2$, where $0 \leq c \leq 1$ and p_1, p_2 are states of $\mathscr{L}$. Since $p^{\sim}(a_n{}^{\perp})=0$,

$$p_1(a_n{}^{\perp}) = p_2(a_n{}^{\perp}) = 0.$$

Hence $p_1(a_n)=p_2(a_n)=1$, and hence the restrictions of p_1 and p_2 to $\mathscr{L}_n$ are states of $\mathscr{L}_n$. Since p is a pure state of $\mathscr{L}_n$, it follows that

$$p(a) = p_1(a) = p_2(a)$$

for all $a<a_n$. But then for any $b \in \mathscr{L}$, we see that $p^{\sim}(b)=p_1(b)=p_2(b)$ since, by (36), $b=(b \wedge a_n) \vee (b \wedge a_n{}^{\perp})$ and $p^{\sim}$, p_1, p_2 vanish for $b \wedge a_n{}^{\perp}$. $p^{\sim}$ is thus a pure state of $\mathscr{L}$. If $m, n \in D$ and $m \neq n$, $\mathscr{P}_m \cap \mathscr{P}_n=\varnothing$, for, if $q \in \mathscr{P}_m$ and $q' \in \mathscr{P}_n$, $q(a_m)=1$ while $q'(a_m)=0$.

We show now that each element of $\mathscr{P}$ lies in some $\mathscr{P}_n$. Suppose that p is a pure state of $\mathscr{L}$. We claim that $p(a_n)=1$ for some $n \in D$. Otherwise there will be some $m \in D$ such that $0<p(a_m)<1$. Define q_1 and q_2 on $\mathscr{L}$ by

$$\begin{aligned} q_1(a) &= (p(a_m))^{-1}p(a \wedge a_m), \\ q_2(a) &= (1-p(a_m))^{-1}p(a \wedge a_m{}^{\perp}). \end{aligned} \tag{55}$$

Since $a_m \in \mathscr{C}$, it can be easily verified using (35) and (36) that q_1 and q_2 are states of $\mathscr{L}$. Since $q_1(a_m)=1$, and $q_2(a_m)=0$, $q_1 \neq q_2$. As

$$p = p(a_m)\cdot q_1+(1-p(a_m))\cdot q_2,$$

p is not pure, a contradiction. Therefore $p(a_n)=1$ for some $n \in D$. The argument given in the previous paragraph can now be repeated to enable us to conclude that the restriction of p to $\mathscr{L}_n$, say q, is a pure state of $\mathscr{L}_n$ and that $p=q^{\sim}$. This proves that $p \in \mathscr{P}_n$ and completes the proof of (54).

We now come to the proof of the last assertion. Assume that $\mathscr{C}$ is continuous and separable. Suppose that p is a state of $\mathscr{L}$. We assert first that for some $c \in \mathscr{C}$, $0<p(c)<1$. If this is not true, then $p(c)=0$ or 1 for each $c \in \mathscr{C}$. Since $\mathscr{C}$ is separable, there exists a countable set

$$\mathscr{D} = \{c_1, c_2, \cdots\}$$

such that $\mathscr{D}$ generates $\mathscr{C}$. Since $p(c_n)=0$ or 1, we may, by replacing c_n by

$c_n^{\perp}$ if necessary, assume that $p(c_n)=1$ for all n. Let $c_0=\bigwedge_n c_n$. Clearly, $p(c_0)=1$. In particular, $c_0\neq 0$. We assert that c_0 is an atom of $\mathscr{C}$. In fact, the subset of all elements c in $\mathscr{C}$ such that either $c\perp c_0$ or $c>c_0$, is a sub-σ-algebra of $\mathscr{C}$ containing all the c_n and so coincides with $\mathscr{C}$; this shows that if $a<c_0$ and $a\in\mathscr{C}$, then $a=0$ or $a=c_0$. This contradicts the assumption that $\mathscr{C}$ is continuous. Consequently, our assumption that p is two-valued on $\mathscr{C}$ is untenable. Let then $d\in\mathscr{C}$ be such that $0<p(d)<1$. Define q_1 and q_2 on $\mathscr{L}$ by

$$q_1(a) = (p(d))^{-1}p(a \wedge d), \qquad q_2(a) = (1-p(d))^{-1}p(a \wedge d^{\perp}).$$

Then $p=p(d)q_1+(1-p(d))q_2$ and p is not pure. The proof of the theorem is complete.

We shall end this section with a few remarks of a general nature regarding the Heisenberg uncertainty principle. We must point out that a considerable part of its meaning is physical and we do not go into this here (cf. the discussions of von Neumann [1] and Heisenberg [1]). We shall be concerned only with mathematical formulations.

Suppose that $\mathfrak{S}$ is a quantum mechanical system and $\mathscr{L}=\mathscr{L}(\mathfrak{S})$ its logic. For any observable x associated with $\mathscr{L}$ and any state p of $\mathscr{L}$, we write $\sigma_p{}^2(x)$ for the variance of x in the state p (defined only when $\mathscr{E}(x^2|p)<\infty$):

$$\sigma_p{}^2(x) = \mathscr{E}(x^2|p)-(\mathscr{E}(x|p))^2. \tag{56}$$

If $\sigma_p(x)=0$, then the probability distribution of x in the state p is concentrated at the point $\mathscr{E}(x|p)$. x has then a sharply defined value in the state p. One might ask the question whether there are states in which *every observable* has, if not a sharply defined value, at least a small variance. The Heisenberg principle of uncertainty asserts that even this is not possible. Roughly speaking, it asserts that in any state, no matter how prepared, there are observables whose distributions have large variances. Notice that this is not possible if the logic is a Boolean σ-algebra of subsets of a space. In fact, let $\mathscr{L}$ be the σ-algebra of subsets of a set X. Then, for any observable with corresponding function f, the observable has a sharply defined value $f(x_0)$ whenever the state p is a measure concentrated in a point x_0 of X. On the other hand, let us assume that $\mathfrak{S}$ is the physical system of some atomic particle and that x and y are the position and momentum observables along some direction in space. If we assume that the logic $\mathscr{L}$ of $\mathfrak{S}$ is the standard logic associated with the complex number field, then

$$\sigma_p(x)\sigma_p(y) \geq \tfrac{1}{2}\hbar, \tag{57}$$

where $\hbar$ is Planck's constant divided by 2π (cf. H. Weyl [1] pp. 77 and 393–394 for a rigorous derivation). From (57) it follows quite simply that in any state in which the distribution of x has small variance (i.e., x is "localized"), the distribution of y has very large variance. It is obvious that this is a property of the structure of the logic $\mathscr{L}$ of $\mathfrak{S}$.

8. AUTOMORPHISMS

Let $\mathscr{L}$ be a logic, $\mathcal{O}$ the set of observables of $\mathscr{L}$, and $\mathscr{S}$ the set of all states of $\mathscr{L}$. The set of all automorphisms of $\mathscr{L}$ forms a group in a natural fashion, denoted by Aut($\mathscr{L}$). We shall now introduce the notion of an automorphism of $\mathscr{S}$. By a *convex automorphism of* $\mathscr{S}$ we mean a map

$$\xi : p \rightarrow p^{\xi} \qquad (p \in \mathscr{S}) \tag{58}$$

such that

(59) (i) ξ is one-one and maps $\mathscr{S}$ onto itself,
(ii) if $c_1, c_2, \cdots$ are numbers ≥ 0 with $\sum_n c_n = 1$, and $p_1, p_2, \cdots \in \mathscr{S}$,

$$\left(\sum_n c_n p_n\right)^{\xi} = \sum_n c_n p_n{}^{\xi}.$$

From (ii) of (59) it follows at once that convex automorphisms of $\mathscr{S}$ send pure states of $\mathscr{L}$ into pure states of $\mathscr{L}$ and that the correspondences induced in the set of pure states are one-one and onto.

If

$$\alpha : a \rightarrow a^{\alpha} \qquad (a \in \mathscr{L}) \tag{60}$$

is an automorphism of $\mathscr{L}$ and if, for $p \in \mathscr{S}$ we define p^{α} by

$$p^{\alpha}(a) = p(a^{(\alpha^{-1})}), \tag{61}$$

then $p \rightarrow p^{\alpha}$ is a convex automorphism of $\mathscr{S}$. We shall call it the automorphism *induced* by α. The set of all convex automorphisms of $\mathscr{S}$ also forms a group denoted by Aut($\mathscr{S}$).

Let $\mathscr{L}$ be a σ-algebra of subsets of a space X, and let

$$t : x \rightarrow t(x) \qquad (x \in X)$$

be any one-one map of X onto itself such that for any set $A \subseteq X$, $A \in \mathscr{L}$ if and only if $t^{-1}(A) \in \mathscr{L}$. When X and $\mathscr{L}$ satisfy some technical regularity conditions, it can be shown that every sufficiently well-behaved automorphism of $\mathscr{S}$ is induced by a mapping t such as the one described above. Suppose X is the phase space of a classical mechanical system with a C^{∞} manifold M as its configuration space. X thus appears (cf. Chapter I) as the cotangent bundle of M. In this case, the subgroup of those automorphisms of $\mathscr{L}$ induced by (both ways) differentiable homeomorphisms of X onto itself and which preserve the canonical 2-form on X, plays undoubtedly a very crucial role (contact transformations).

The axioms satisfied by an arbitrary logic are too weak to allow any reasonable determination of its automorphisms. In the next chapter we shall determine the automorphisms associated with standard logics. In the meantime, we shall briefly illustrate with two (among many) examples the important role played by the group Aut($\mathscr{S}$) in physical problems.

As a first example we consider the problem of describing the dynamics of a physical system $\mathfrak{S}$. Let $\mathscr{L}$ be the logic of $\mathfrak{S}$, $\mathscr{S}$ the convex set of all its states. Then, for each real number t, there exists a unique one-one map of $\mathscr{S}$ onto itself, say $D(t)$, with the following physical interpretation; if $p \in \mathscr{S}$ is the state of $\mathfrak{S}$ at time t_0, then $D(t)p$ is the state of $\mathfrak{S}$ at time $t+t_0$. The dynamical group $\mathscr{D}$ where

$$\mathscr{D} : t \to D(t) \tag{62}$$

satisfies the conditions (3) of Chapter I. We now postulate that each $D(t)$ is in fact an automorphism of $\mathscr{S}$, i.e., if $p_1, p_2, \cdots$ is a sequence of elements of $\mathscr{S}$ and $c_1, c_2, \cdots$ a sequence of nonnegative numbers such that $c_1+c_2+\cdots=1$, then, for each t,

$$D(t)\Big(\sum_i c_i p_i\Big) = \sum_i c_i D(t)p_i. \tag{63}$$

This postulate may be justified by the following argument. Suppose one does not know the state at time 0 but only that it is one of $p_1, p_2, \cdots$ with probabilities $c_1, c_2, \cdots$, respectively. Then from the fact that the state at time t is $D(t)p_i$ if the state time 0 is p_i, it follows that at time t the state of the system $\mathfrak{S}$ is one of $D(t)p_1$, $D(t)p_2, \cdots$ with the same probabilities $c_1, c_2, \cdots$. But the statement that the state of $\mathscr{L}$ (or $\mathfrak{S}$) is one of $q_1, q_2, \cdots$ with probabilities $c_1, c_2, \cdots$ is physically equivalent to the statement that the state of $\mathscr{L}$ is $c_1q_1+c_2q_2+\cdots$. This proves the validity of (63). It follows now that each $D(t)$ leaves invariant the subset of pure states of $\mathscr{L}$ and induces a one-one mapping on it. Furthermore, if $a \in \mathscr{L}$, then, for each $p \in \mathscr{S}$,

$$t \to (D(t)p)(a) \tag{64}$$

is a real valued function of t; it is physically reasonable to assume that for any $\mathfrak{S}$, this function is Borel for each $p \in \mathscr{S}$ and $a \in \mathscr{L}$. If this condition is satisfied, we call $\mathscr{D}$ a *Borel one-parameter group. Thus the dynamical group of $\mathfrak{S}$ may be assumed to be a Borel one-parameter group of convex automorphisms of $\mathscr{S}$.*

This leads us to introduce the concept of representations of groups. Let G be a locally compact topological group satisfying the second axiom of countability. The sets of the smallest σ algebra of subsets of G

containing the open sets are called the *Borel sets* in G. By a *representation of G in* $\mathrm{Aut}(\mathscr{S})$ we mean a map

(65) $$T : g \to T_g \qquad (g \in G)$$

of G such that

(66)
- (i) each $T_g \in \mathrm{Aut}(\mathscr{S})$,
- (ii) $T_{g_1 g_2} = T_{g_1} T_{g_2} \qquad (g_1, g_2 \in G)$,
- (iii) for each $a \in \mathscr{L}$ and $p \in \mathscr{S}$, the function $g \to (T_g p)(a)$ is Borel on G.

With these definitions we may say that the dynamical group of a system $\mathfrak{S}$ is determined by a representation of the additive group of real numbers into $\mathrm{Aut}(\mathscr{S})$, where $\mathscr{S}$ is the set of states of the logic $\mathscr{L}$ of $\mathfrak{S}$.

As our second example, we consider the problem of describing the quantum mechanical theory of a free particle. The configuration space of the particle in classical mechanics is the three-dimensional space, denoted by M. Let $\mathscr{L}$ be the logic of the physical system under consideration and $\mathscr{S}$ the convex set of its states. Suppose now an observer O wants to describe the system. He would then select a point of M, an orthogonal frame at O, and map M onto R^3 by a map

(67) $$\gamma_O : M \to R^3.$$

Let O' be another observer and let $\gamma_{O'}$ be the corresponding map. Then there exists a unique Euclidean motion $D_{O,O'}$ of R^3 such that for all $m \in M$

(68) $$\gamma_{O'}(m) = D_{O,O'}(\gamma_O(m)).$$

Suppose now O describes the state of the system $\mathfrak{S}$ by an element $p_O \in \mathscr{S}$. Then, O', using his orientation, would describe the same physical state by an element $p_{O'}$ of $\mathscr{S}$. It is clear that the map $p_O \to p_{O'}$ is an *automorphism*, say $\alpha_{O,O'}$ of $\mathscr{S}$. In other words, $\alpha_{O,O'}$ has the following physical interpretation: if the physical state of $\mathfrak{S}$ is described by O by the element $g \in \mathscr{S}$, then $\alpha_{O,O'}(g)$ is the element of $\mathscr{S}$ describing the same physical state of $\mathfrak{S}$ from the point of view of O'. It is obvious that $\alpha_{O,O'}$ is the identity for $O = O'$, and

(69) $$\alpha_{O,O''} = \alpha_{O',O''}\alpha_{O,O'}.$$

We now remark that the basic physical laws of $\mathfrak{S}$ are independent of the choice of observers and their frames. Therefore, we may conclude that $\alpha_{O,O'}$ depends only on the transformation $D_{O,O'}$. Let G be the group of all

Euclidean motions of R^3. Then, from (68), (69), and the above remarks on the $\alpha_{O,O'}$, we infer the existence of a homomorphism

$$T : g \to T_g \tag{70}$$

of G into $\mathrm{Aut}(\mathscr{S})$ with the property that whenever O and O' are two observers of $\mathfrak{S}$,

$$\alpha_{O,O'} = T_{D_{O,O'}}. \tag{71}$$

We may reasonably make the assumption that for each $a \in \mathscr{L}$ and $p \in \mathscr{S}$, $g \to (T_g p)(a)$ is a Borel function on G. T is thus a representation of G in $\mathrm{Aut}(\mathscr{S})$.

We see thus that the description of $\mathfrak{S}$ gives rise to a representation of G in $\mathrm{Aut}(\mathscr{S})$. The determination of such representations is therefore the basic problem to be solved in the mathematical description of the physical system of a free particle.

The reader might notice that the arguments given above are very general. In fact, the same type of arguments lead to the remarkable fact that with every *relativistic free particle* there is associated a representation of the *inhomogeneous Lorentz group*. These representations were first completely determined by Wigner under the assumption that the logic of $\mathfrak{S}$ was standard. This led in turn to his celebrated classification of the relativistic wave equations [1].

NOTES ON CHAPTER III

1. As a general reference on lattice theory we refer to the book of F. Maeda and S. Maeda, *Theory of Symmetric Lattices*, Springer-Verlag, Berlin, 1970.

2. There has been a good amount of work by many people in recent years on the structure of logics and probability measures on them, motivated mostly by the relevance of such problems to the foundations of Quantum Mechanics. The encyclopedia volume of E. G. Beltrametti and G. Cassinelli (*The Logic of Quantum Mechanics*, Volume 15, *Encyclopedia of Mathematics and Its Applications*, Addison-Wesley, Reading, Mass., 1981) gives an excellent overview of this work (Chs. 10–20). In addition, this book presents a well-organized, comprehensive discussion of the foundations of Quantum Mechanics from different points of view with many references to current literature, examples, and exercises. Especially illuminating are the critiques of controversial topics such as hidden variables (Chs. 15, 25) consistency of physical interpretation (Chs. 7, 8), the process of measurement (Ch. 26), and so on. As additional references we mention the following:

J. M. Jauch, *Foundations of Quantum Mechanics*, Addison-Wesley, Reading, Mass., 1968.

C. Piron, *Foundations of Quantum Physics*, Benjamin, Reading, Mass., 1976.
M. Jammer, *The Philosophy of Quantum Mechanics*, Wiley, New York, 1974.
C. A. Hooker (Ed.), *The Logico-algebraic Approach to Quantum Mechanics*, Vol. 1, Reidel, Dordrecht, 1975.
E. G. Beltrametti and B. C. van Frassen (Eds.), *Current Issues in Quantum Logic*, Plenum, New York, 1981.

3. The structure of the states of orthocomplemented continuous geometries is still mysterious. One of the most striking results is due to von Neumann who showed that in the presence of a transition probability function with suitable invariance properties, such a lattice is nothing but the lattice of projections of a type II_1 factor (Continuous geometries with a transition probability, Mem. Amer. Math. Soc. No. 252, Providence, R.I., 1981).

4. It is possible to determine, in the standard model, the states for which the lower bound of the Heisenberg uncertainty relations (57) is attained, for the quantum mechanical system of a one-dimensional particle. Let $\mathscr{H} = L^2(-\infty,\infty)$,

$$q = \text{operator of multiplication by } x, p = \frac{\hbar}{i}\frac{d}{dx};$$

then

$$\sigma_\varphi(q)\cdot\sigma_\varphi(p) = \frac{\hbar}{2}$$

for the family of states $\varphi = \varphi_{\gamma,\sigma,\rho}$ $(\gamma,\sigma,\rho\in\mathbf{R}, \gamma>0)$ given by

$$\varphi(x) = \left(\frac{\gamma}{\hbar\pi}\right)^{\frac{1}{4}} \exp\{-(\gamma/2\hbar)(x-\sigma)^2 + (p/\hbar)ix\};$$

then

$$\sigma_\varphi(q) = (\hbar/2\gamma)^{\frac{1}{2}}, \qquad \sigma_\varphi(p) = (\hbar\gamma/2)^{\frac{1}{2}}$$

(cf. von Neumann [1], pp. 233–237).

As we shall see in greater detail in Chapter IV, any vector φ of unit norm in $L^2(-\infty,\infty)$ defines a pure state of our system; the probability for $S\in\mathscr{L}(\mathscr{H})$ is $\|P^S\varphi\|^2$, P being the orthogonal projection $\mathscr{H}\to S\cdot\varphi$ is called the *wave function*, and since the spectral measure of q is $E\to q_E$ where q_E is multiplication by the characteristic function of E, the probability distribution of q in the state φ is the measure

$$\mu : E \to \int_E |\varphi|^2\,dx.$$

Thus the wave function allows one to calculate, by this formula, for any region $E\subset(-\infty,\infty)$, the probability $\mu(E)$ of finding the particle inside E. This is the simplest instance of the *statistical interpretation* of Quantum Mechanics which is due to Max Born (Z. Physik, **37** (1926), pp. 863–867).

CHAPTER IV

LOGICS ASSOCIATED WITH HILBERT SPACES

1. THE LATTICE OF SUBSPACES OF A BANACH SPACE

The general theory that was developed in Chapter III is concerned mainly with the broad principles of Quantum Theory. Further development of the theory, such as a discussion of simple quantum mechanical systems, leads to problems of a more technical nature. These problems are, however, difficult to answer in the context of abstract logics, and therefore, in dealing with them, it becomes necessary to restrict the class of logics under consideration. In this chapter we shall examine some of these logics and describe, in a somewhat more concrete fashion than in Chapter III, the observables, states, and symmetries associated with them.

We have seen in Chapter II that a large class of orthocomplemented geometries are obtained by considering Hilbert space structures on finite dimensional vector spaces. These logics are, however, not of much use in discussing even very simple atomic systems, since every observable associated with them is finite. Consequently, if one wants to obtain realistic examples, one is forced to deal with *infinite dimensional* vector spaces.

It is known (cf. Baer [1]) that the lattice of *all* linear manifolds of an infinite dimensional vector space does not admit any polarity. Therefore it becomes necessary to consider lattices $\mathscr{L}$ such that the elements of $\mathscr{L}$ are certain linear manifolds of a given vector space V and, wherein, the partial ordering is the inclusion relation. It must be noted that the lattice sum of two elements of such a lattice contains, but in general does not coincide with, the algebraic sum in V. A simple and interesting class of such lattices may be obtained in the following way. Let $\mathbf{D}$ be a topological division ring, say, for example, one of $\mathbf{R}$, $\mathbf{C}$, or $\mathbf{H}$. Let V be a topological vector space over $\mathbf{D}$ (cf. Bourbaki [2] for the definitions and elementary properties). Then, the set of all *closed* linear manifolds of V is a lattice—in fact, a complete lattice. One may then construct examples of logics whose underlying lattices are these lattices of closed manifolds. The description of such logics would provide us with a large class of examples of logics which are complex enough to serve as models for the logics of atomic systems.

The main purpose of this section is to prove a theorem of Kakutani and Mackey ([1]), which asserts that if V is a Banach space, the above procedure leads to no new examples of logics other than the standard ones. In particular, if the underlying lattice of a logic is isomorphic to the lattice of all closed linear manifolds of a separable *infinite dimensional* Banach space over one of **R**, **C**, or **H**, then the logic is standard.

Let **D** be one of **R**, **C**, or **H** and V any vector space over **D**. We shall say that a topology $\mathscr{T}$ on V is *compatible* with the vector space structure of V if it is Hausdorff, and, if the map, $a, b, x, y \rightarrow a \cdot x + b \cdot y$ of $\mathbf{D} \times \mathbf{D} \times V \times V$ into V, is continuous. A *topological vector space* over **D** is a pair $(V, \mathscr{T})$, where V is a vector space over **D**, and $\mathscr{T}$, a topology over V compatible with its vector space structure. By the usual abuse of language we shall say that V is a topological vector space. We shall denote by $\mathscr{L}(V,\mathbf{D})$ the lattice of all closed linear manifolds of V. This is a complete lattice. If we denote by $\vee$ and $\wedge$ the lattice operations in $\mathscr{L}(V,\mathbf{D})$, then, for any family $\{M_\alpha\}$ of closed linear manifolds of V, we have:

$$\bigwedge_\alpha M_\alpha = \bigcap_\alpha M_\alpha, \tag{1}$$

$$\bigvee_\alpha M_\alpha = \Big(\sum_\alpha M_\alpha\Big)^-, \tag{2}$$

where $\sum$ denotes the algebraic sum, and the bar denotes closure. If N is any *finite dimensional* linear manifold of V, then N is necessarily closed and for any closed linear manifold M, $N \vee M = N + M$. Moreover, if M is a closed linear manifold whose codimension $(= \dim V/M)$ is finite and N any algebraic complement of M, i.e., such that $M \cap N = 0$, $M + N = V$, then N is necessarily closed and $V = M \vee N$ (for these facts cf. Bourbaki [2], Ch. 1).

We shall recall the notion of Banach spaces. Let **D** be one of **R**, **C**, or **H** and V a vector space over **D**. A *norm* over V is a function $x \rightarrow \|x\|$ of V into the real numbers such that

$$\begin{aligned} &\text{(i)} && \|x\| \geq 0, = 0 \quad \text{if and only if} \quad x = 0, \\ &\text{(ii)} && \|x+y\| \leq \|x\| + \|y\| \quad \text{for} \quad x, y \in V, \\ &\text{(iii)} && \|cx\| = |c| \cdot \|x\| \quad \text{for} \quad c \in \mathbf{D}, x \in V. \end{aligned} \tag{3}$$

A Banach space over **D** is a pair $(V, \|\cdot\|)$, where V is a vector space over **D** and $\|\cdot\|$ a norm on V such that V is complete under the metric d defined by $d(x,y) = \|x-y\|$. A Banach space over **D** is obviously a topological vector space over **D**.

We are now in a position to formulate the theorem of Kakutani and Mackey. We recall the notion of Hilbert spaces and inner products associated with **D** (cf. Chapter II, section 4).

Theorem 4.1 (*Kakutani-Mackey* [1]). *Let* $\mathbf{D}$ *be one of the three division rings* $\mathbf{R}$, $\mathbf{C}$, *or* $\mathbf{H}$, V *an infinite dimensional Banach space over* $\mathbf{D}$, *and* $\mathscr{L}=\mathscr{L}(V,\mathbf{D})$ *the lattice of all closed linear manifolds of* V. *Suppose that* $\perp(M \rightarrow M^{\perp})$ *is an orthocomplementation of* $\mathscr{L}$. *Then there exists a* $\mathbf{D}$-*valued inner product* $\langle . , . \rangle$ *on* $V \times V$ *such that* (i) V *becomes, under* $\langle . , . \rangle$, *a Hilbert space over* $\mathbf{D}$; (ii) *the topology of* V, *induced by the norm associated with* $\langle . , . \rangle$, *coincides with its original topology; and* (iii) *the map* $M \rightarrow M^{\perp}$ *coincides with the orthocomplementation induced by* $\langle . , . \rangle$. *In particular, if the underlying lattice of a logic is isomorphic to the lattice of closed linear manifolds of a separable Banach space of infinite dimension over* $\mathbf{D}$, *then the logic is standard.*

The proof is rather long and depends on a series of lemmas. With later applications in mind we formulate our first lemma somewhat more generally than is necessary for our immediate purposes.

Lemma 4.2. *Let* $\mathbf{K}$ *be a division ring and* W *a vector space of infinite dimension over* $\mathbf{K}$. *Let* $\mathscr{M}$ *be a lattice of linear manifolds of* W *such that* (i) $\mathscr{M}$ *contains all finite dimensional linear manifolds of* W, *and* (ii) *if* $M, N \in \mathscr{M}$ *and at least one of them is finite dimensional, then* $M \vee N = M + N$. *Suppose* $\mathscr{M}(M \rightarrow M^{\perp})$ *is an orthocomplementation in* $\mathscr{M}$. *Let* $w_0 \in W$ *be any nonzero vector. Then there exists an involutive anti-automorphism* θ *of* $\mathbf{K}$ *and a symmetric* θ-*bilinear form* $\langle . , . \rangle$ *on* $W \times W$ *such that* (i) $\langle w_0, w_0 \rangle = 1$ *and* (ii) $\langle x,y \rangle = 0$ *if and only if* $x \in (\mathbf{K}\cdot y)^{\perp} \cdot \theta$ *and* $\langle . , . \rangle$ *are uniquely determined. The form* $\langle . , . \rangle$ *is definite, i.e.,* $\langle x,x \rangle = 0$ *if and only if* $x = 0$.

Proof. Let $F \subseteq W$ be a finite dimensional linear manifold. Then $F \in \mathscr{M}$. For any linear manifold $L \subseteq F$ we have $L \in \mathscr{M}$ also. We now define

$$\xi_F(L) = L^{\perp} \cap F. \tag{4}$$

Note that if M is finite dimensional and $N \in \mathscr{M}$, then $M \wedge N = M \cap N$. In fact, $M \wedge N \subseteq M \cap N$ always while the finite dimensionality of $M \cap N$ implies, by (i), that $M \cap N \in \mathscr{M}$, so that $M \cap N \subseteq M \wedge N$.

Let $\mathscr{M}_F$ be the projective geometry of all linear manifolds $L \subseteq F$. ξ_F is a map of $\mathscr{M}_F$ into itself. We shall show first that ξ_F is an orthocomplementation of $\mathscr{M}_F$. Clearly $\xi_F(0) = F$, $\xi_F(F) = F^{\perp} \cap F = F^{\perp} \wedge F = 0$. Since $M \rightarrow M^{\perp}$ is order inverting, it is obvious that ξ_F is order inverting. We observe next that for $L \in \mathscr{M}_F$, $L \cap \xi_F(L) \subseteq L \cap L^{\perp} = L \wedge L^{\perp} = 0$. On the other hand, $W = L \vee L^{\perp} = L + L^{\perp}$ so that $F = L + \xi_F(L)$. Thus L and $\xi_F(L)$ are complements of each other in $\mathscr{M}_F$. To complete the proof that ξ_F is an orthocomplementation, it remains to prove that ξ_F is involutive. For $L \in \mathscr{M}_F$, $\xi_F(L) \subseteq L^{\perp}$ so that $L \subseteq \xi_F(L)^{\perp}$. This shows that $L \subseteq \xi_F(L)^{\perp} \cap F = \xi_F(\xi_F(L))$ for all $L \in \mathscr{M}_F$. Let $L' = \xi_F(L)$. Then, by what we proved just now, $\xi_F(L')$ is a complement of L' in $\mathscr{M}_F$. But L is also a complement of $L' = \xi_F(L)$ and we have seen that $L \subseteq \xi_F(L')$. This shows that L must coincide with $\xi_F(L')$. We see thus that ξ_F is an orthocomplementation of $\mathscr{M}_F$.

We shall fix a nonzero vector $w_0 \in W$. Let $\mathscr{F}$ be the class of all finite dimensional linear manifolds $F \subseteq W$ such that (i) $w_0 \in F$, (ii) $\dim F \geq 3$. By theorem 2.7 there exists, for any $F \in \mathscr{F}$, an involutive anti-automorphism θ_F of **K** and a nonsingular symmetric θ_F-bilinear form $\langle . , . \rangle_F$ on $F \times F$ such that $\langle . , . \rangle_F$ induces ξ_F; i.e., for L,

$$\xi_F(L) = \{u : u \in F, \langle x,u \rangle_F = 0 \quad \text{for} \quad \text{all } x \in L\} \tag{5}$$

and

$$\langle w_0, w_0 \rangle = 1. \tag{6}$$

Moreover, (6) determines θ_F and $\langle . , . \rangle_F$ uniquely. Suppose now that F_1, $F_2 \in \mathscr{F}$ with $F_1 \subseteq F_2$. Then, for any $L \in \mathscr{M}_{F_1}$,

$$\xi_{F_1}(L) = L^{\perp} \cap F_1 = \{\xi_{F_2}(L)\} \cap F_1.$$

This shows that for any $L \in \mathscr{M}_{F_1}$,

$$\xi_{F_1}(L) = \{u : u \in F_1, \langle x,u \rangle_{F_2} = 0 \quad \text{for all} \quad x \in L\}. \tag{7}$$

In other words, the restriction of $\langle . , . \rangle_{F_2}$ to $F_1 \times F_1$ is a θ_{F_2}-bilinear form which also induces the orthocomplementation ξ_{F_1}. In view of the normalizing condition (6), we may conclude at once that

$$\begin{aligned} \theta_{F_2} &= \theta_{F_1}, \\ \langle x,y \rangle_{F_2} &= \langle x,y \rangle_{F_1} \qquad (x, y \in F_1). \end{aligned} \tag{8}$$

Since any two elements of $\mathscr{F}$ are contained in a third element of $\mathscr{F}$, we easily deduce from (8) that there exists an involutive anti-automorphism θ of **K**, and a symmetric θ-bilinear form $\langle . , . \rangle$ on $W \times W$ such that (i) $\theta = \theta_F$ for all $F \in \mathscr{F}$, and (ii) the restriction of $\langle . , . \rangle$ to $F \times F$ is $\langle . , . \rangle_F$ for all $F \in \mathscr{F}$. Lemma 4.2 follows at once from this.

We now resume our discussion of the lattice $\mathscr{L}$ of all closed linear manifolds of the Banach space V over **D**. We take $V = W$ in lemma 4.2 and obtain an involutive anti-automorphism θ of **D** and a definite symmetric θ-bilinear form $\langle . , . \rangle$ on $V \times V$ such that $\langle x,y \rangle = 0$ if and only if $x \perp y$. We fix a nonzero vector $v_0 \in V$; then θ and $\langle . , . \rangle$ are uniquely determined by the condition $\langle v_0, v_0 \rangle = 1$.

Let V^* be the set of all *continuous* linear maps of V into **D**. Then V^* is a vector space over $\mathbf{D}^0$, the division ring dual to **D**. For $u \in V$, let β_u be the linear map of V into **D** defined by

$$\beta_u(x) = \langle x,u \rangle \qquad (x \in V). \tag{9}$$

We shall also use the conjugate norm on V^*; this is defined by

$$\|\lambda\| = \sup_{\|u\| \leq 1} |\lambda(u)| \qquad (\lambda \in V^*). \tag{10}$$

It is well known that V^*, under this norm, is a Banach space over the division ring $\mathbf{D}^0$.

Lemma 4.3. *$\beta_u \in V^*$ for any $u \in V$. Moreover, $u \to \beta_u$ is an additive isomorphism of the abelian group V onto the abelian group V^*.*

Proof. It is a well known fact (cf. Bourbaki [2]) that a linear map λ of V into $\mathbf{D}$ is continuous if and only if the linear manifold

$$Z_\lambda = \{x : x \in V, \quad \lambda(x) = 0\} \tag{11}$$

is closed in V. Let now $\lambda = \beta_u$, and write Z_u for Z_λ, $u \in V$. If $y \in V$, $y \in Z_u$ if and only if $\langle y,u \rangle = 0$, i.e., if and only if $y \perp u$. Hence

$$Z_u = (\mathbf{D} \cdot u)^\perp. \tag{12}$$

But $(\mathbf{D} \cdot u)^\perp \in \mathscr{L}$ and is therefore closed by assumption. This proves that $\beta_u \in V^*$. The map $u \to \beta_u$ is clearly additive. If, for $u \in V$, $\beta_u = 0$, then $\langle u,u \rangle = \beta_u(u) = 0$; this implies that $u = 0$. In order to complete the proof of the lemma it remains to prove only that $u \to \beta_u$ maps V onto V^*. Suppose $\lambda \in V^*$ and $\neq 0$. Let Z_λ be defined by (11). Then $Z_\lambda \in \mathscr{L}$ and $Z_\lambda \neq V$. Therefore $Z_\lambda{}^\perp \neq 0$. Let u be a nonzero vector in $Z_\lambda{}^\perp$. Since $\langle x,u \rangle = 0$ for all $x \in Z_\lambda$, β_u vanishes over Z_λ; and since $u \neq 0$, $\beta_u \neq 0$. Since Z_λ and Z_u are both of codimension 1 and $Z_\lambda \subseteq Z_u$, we infer that $Z_\lambda = Z_u$. Let $x_0 \in V$ be such that $\lambda(x_0) = 1$. Since $x_0 \notin Z_\lambda$, $x_0 \notin Z_u$ and hence

$$\beta_u(x_0) = \langle x_0,u \rangle \neq 0.$$

Let $c = \langle x_0,u \rangle^{-1}$ and let us write $u_0 = c^\theta u$. Then for $x \in V$,

$$\langle x,u_0 \rangle = \langle x,u \rangle c.$$

This shows that $\beta_{u_0}(x_0) = 1 = \lambda(x_0)$ and implies that $\beta_{u_0} = \lambda$.

The main question at this stage is whether θ is continuous. The nontrivial case is when $\mathbf{D} = \mathbf{C}$ and we handle it first. It is at this point that essential use is made of the infinite dimensionality of V.

Lemma 4.4. *Let W be a complex Banach space of infinite dimension. Then one can find a sequence of vectors $x_1, x_2, \cdots \in W$ with the property that, for any bounded infinite sequence $\{c_n\}$ of complex numbers, there exists a continuous linear functional λ on W such that*

$$\lambda(x_n) = c_n \qquad (n = 1, 2, \cdots). \tag{13}$$

Proof. Let $\|\cdot\|$ be the norm on W. Let W^* be the space of all complex valued continuous linear functionals on W. We define $x_1 \in W$ to be an arbitrary vector of norm 1 and choose $x_1{}^* \in W^*$ such that

$$\|x_1{}^*\| = 1,$$
$$x_1{}^*(x_1) = 1.$$

The celebrated Hahn-Banach theorem ensures that such an $x_1{}^* \in W^*$ exists. We shall now define by induction two sequences $\{x_n\}$ and $\{x_n{}^*\}$ in W and W^*, respectively. Suppose $x_1, x_2, \cdots, x_n \in W$ and $x_1{}^*, \cdots, x_n{}^* \in W^*$ have been already constructed so that

$$\text{(14)} \qquad \begin{aligned} x_i{}^*(x_j) &= \delta_{ij} \text{ (Kronecker delta)} \qquad (i, j = 1, 2, \cdots, n), \\ \|x_i{}^*\| &= 2^{-(i-1)} \qquad (i = 1, 2, \cdots, n). \end{aligned}$$

Since W is infinite dimensional, there is a nonzero vector x_{n+1} in W linearly independent of $x_1, \cdots, x_n$ such that

$$x_i{}^*(x_{n+1}) = 0 \qquad (i = 1, 2, \cdots, n).$$

Using the Hahn-Banach theorem once again, we can find $x_{n+1}^* \in W^*$ such that $x_{n+1}^*(x_{n+1}) \neq 0$, $x_{n+1}^*(x_i) = 0$, $i \leq n$. Replacing x_{n+1} and x_{n+1}^* by suitable multiples, we may assume that $\|x_{n+1}^*\| = 2^{-n}$, $x_{n+1}^*(x_{n+1}) = 1$. The sequences $x_1, x_2, \cdots, x_{n+1}$ and $x_1{}^*, \cdots, x_{n+1}^*$ then satisfy (14) with n replaced by $n+1$. By induction it follows that there exist infinite sequences $\{x_n\}$ and $\{x_n{}^*\}$ such that $x_n \in W$ for all n, $x_n{}^* \in W^*$ for all n, and

$$\text{(15)} \qquad \begin{aligned} x_i{}^*(x_j) &= \delta_{ij} \qquad (i, j = 1, 2, \cdots), \\ \|x_i{}^*\| &= 2^{-(i-1)} \qquad (i = 1, 2, \cdots). \end{aligned}$$

We claim that the sequence $\{x_n\}$ has the properties asserted in the lemma. In fact, let $\{c_n\}$ be any bounded sequence of complex numbers. For $x \in W$, let $\lambda(x)$ be defined by

$$\text{(16)} \qquad \lambda(x) = \sum_{n=1}^{\infty} c_n x_n{}^*(x).$$

The inequality $|x_n{}^*(x)| \leq \|x_n{}^*\| \cdot \|x\| = 2^{-(n-1)}\|x\|$ shows that λ is well defined and that $x \to \lambda(x)$ is a continuous linear functional on W, i.e., $\lambda \in W^*$. Clearly, $\lambda(x_n) = c_n$ for all n.

We now resume our discussion of V and θ.

Lemma 4.5. *If* $\mathbf{D} = \mathbf{R}$, θ *is the identity; if* $\mathbf{D} = \mathbf{C}$, θ *is the conjugation; and if* $\mathbf{D} = \mathbf{H}$, θ *is the canonical conjugation.*

Proof. Let F be any linear manifold of finite dimension ≥ 3 containing v_0. The map $L \to L^{\perp} \cap F$ is, as we saw in the proof of lemma 4.2, an orthocomplementation in the geometry of subspaces of F and is, moreover, induced by $\langle . , . \rangle_F$ (the restriction of $\langle . , . \rangle$ to $F \times F$). From theorem 2.8 we deduce that θ is the identity if $\mathbf{D} = \mathbf{R}$ and is the canonical conjugation if $\mathbf{D} = \mathbf{H}$. Suppose now that $\mathbf{D} = \mathbf{C}$. Clearly, as $\langle x,x \rangle_F = 0$ for an $x \in F$ implies that $x = 0$, θ is not the identity. θ will therefore have to be the conjugation in this case, provided we show that it is continuous. Since θ is additive, this reduces to showing that if $c_n \to 0$ in $\mathbf{D}$, $\{c_n{}^\theta\}$ is a bounded sequence. Suppose then that $\mathbf{D} = \mathbf{C}$ and that θ is not the conjugation. Then there exists a sequence $\{c_n\}$ in $\mathbf{D}$ such that

$$\begin{aligned} & c_n \to 0, \\ (17) \qquad & |c_n{}^\theta| \to \infty. \end{aligned}$$

Since $\dim V = \infty$, we can select a sequence $\{x_n\}$ in V such that it has the property described in lemma 4.4. Let β_n be the element $x \to \langle x,x_n \rangle$ of V^* (cf. lemma 4.3). Since $|c_n{}^\theta| \to \infty$, we may, by passing to a subsequence if necessary, assume in addition that for all n,

$$(18) \qquad |c_n{}^\theta| > n \cdot \|\beta_n\|.$$

By lemma 4.4 we can find a $\lambda \in V^*$ such that $\lambda(x_n) = c_n$ for all n. Evidently, $\lambda \neq 0$. Moreover, it follows from lemma 4.3 that there is a $u \neq 0$ in V such that $\lambda = \beta_u$. Then $\langle u,u \rangle \neq 0$ and $\langle x_n,u \rangle = c_n$ for all n. If

$$(19) \qquad z_n = \langle u,u \rangle^{-1} u - c_n{}^{-1} x_n,$$

then

$$\lambda(z_n) = \langle z_n,u \rangle = 0$$

for all n. But then, as $\langle . , . \rangle$ is symmetric,

$$\langle u,z_n \rangle = 0$$

for all n, i.e.,

$$(20) \qquad (\langle u,u \rangle^\theta)^{-1} \langle u,u \rangle - (c_n{}^\theta)^{-1} \langle u,x_n \rangle = 0.$$

Since $|(c_n{}^\theta)^{-1} \langle u,x_n \rangle| \leq |c_n{}^\theta|^{-1} \|\beta_n\| \|u\|$, (18) implies that

$$(c_n{}^\theta)^{-1} \langle u,x_n \rangle \to 0 \qquad (n \to \infty)$$

and hence we infer from (20) that $\langle u,u \rangle (\langle u,u \rangle^\theta)^{-1} = 0$, which is impossible as $u \neq 0$. This contradiction proves that θ must be the conjugation and completes the proof of lemma 4.5.

Lemma 4.6. *$\beta(u \to \beta_u)$ is a continuous map of V onto V^*.*

Proof. Since we may obviously consider $\mathbf{R}$ as a subfield of $\mathbf{D}$ as well as $\mathbf{D}^0$, V and V^* may both be regarded as Banach spaces over $\mathbf{R}$, denoted by $V_{\mathbf{R}}$ and $V_{\mathbf{R}}{}^*$, respectively. From lemma 4.5 we conclude that $\beta(u \to \beta_u)$ is an $\mathbf{R}$-linear map. As β is defined on the whole of $V_{\mathbf{R}}$, we may use the closed graph theorem to reduce the proof of the continuity of β to the proof that β is a *closed* map. Let $\{x_n\}$ be a sequence in $V_{\mathbf{R}}$ such that

$x_n \to x$ and $\beta_{x_n} \to \lambda$. We must prove that $\lambda = \beta_x$. Since β is onto, we may assume that $\lambda = \beta_u$ for some $u \in V$. From the formula (10) for the norm in V^*, it follows that $\beta_{x_n}(v) \to \beta_u(v)$ for all $v \in V$. Since, by lemma 4.5, θ is continuous, we have, for $v \in V$,

$$\begin{aligned} \langle x,v \rangle &= \lim_n \langle x_n,v \rangle \\ &= \lim_n (\langle v,x_n \rangle^\theta) \\ &= \langle v,u \rangle^\theta \\ &= \langle u,v \rangle. \end{aligned}$$

This proves that $x = u$ or $\beta_x = \lambda$. We have thus proved lemma 4.6.

Lemma 4.5 shows that $\langle .\,,. \rangle$ is an inner product on $V \times V$. Thus, on defining $\|z\|_0{}^2 = \langle z,z \rangle$, we obtain a norm on V. V is a pre-Hilbert space with respect to $\langle .\,,. \rangle$.

Lemma 4.7. *V is a Hilbert space under $\langle .\,,. \rangle$.*

Proof. We must prove that V is complete under $\|\cdot\|_0$. Let $\{x_n\}$ be a sequence in V such that $\|x_n - x_m\|_0{}^2 \to 0$, $n, m \to \infty$. Then $\langle y,x_n - x_m \rangle \to 0$ as $n, m \to \infty$ for each $y \in V$. Let $\lambda(y) = \lim_n \langle y,x_n \rangle$. Then λ is a linear map of V into $\mathbf{D}$. Since $\beta_{x_n} \in V_{\mathbf{R}}{}^*$ for all n, we may apply the well known uniform boundedness principle to $V_{\mathbf{R}}$, a Banach space over $\mathbf{R}$, to conclude that $\lambda \in V^*$. By lemma 4.3 there exists a $z \in V$ such that $\lambda = \beta_z$. We claim that $\|x_n - z\|_0 \to 0$. Let $\varepsilon > 0$ be arbitrary and N be such that

$$\|x_n - x_m\|_0 \leq \varepsilon$$

for $n, m \geq N$. Then, for $y \in V$,

$$|\langle y, x_n - x_m \rangle| \leq \varepsilon \|y\|_0 \qquad (n, m \geq N)$$

and hence

$$|\langle y, x_n - z \rangle| \leq \varepsilon \|y\|_0 \qquad (n \geq N),$$

which shows that $\|x_n - z\|_0 \leq \varepsilon$ for $n \geq N$.

Lemma 4.8. *The topologies induced on V by $\|\cdot\|$ (the original norm) and $\|\cdot\|_0$ coincide.*

Proof. Let $y, x_1, x_2, \cdots$ be elements in V. We must show that as $n \to \infty$, $\|x_n - y\|_0 \to 0$ if and only if $\|x_n - y\| \to 0$. Since V is a Banach space over $\mathbf{R}$ under both $\|\cdot\|$ and $\|\cdot\|_0$, it is sufficient, in view of the open mapping theorem, to show the implication in one direction only. Let then $\|x_n - y\| \to 0$. Since the map β from V to V^* is continuous, $\beta_{x_n} - \beta_y \to 0$ in V^* so that

$$\sup_{\|z\| \leq 1} |\langle z, x_n - y \rangle| \to 0 \qquad (n \to \infty). \tag{21}$$

Now, $\|x_n - y\| \leq 1$ for all sufficiently large n and hence we infer from (21) that

$$\|x_n - y\|_0^2 = \langle x_n - y, x_n - y\rangle \to 0,$$

which is the relation required.

We can now complete the proof of theorem 4.1. V is a Hilbert space under $\langle . , . \rangle$ and the topology induced by the corresponding norm has been shown to coincide with the original topology for V. Consequently, the same linear manifolds are closed in both topologies. Let $\xi(M \to \xi(M))$ be the orthocomplementation of the Hilbert space V. Now, $\langle y,x\rangle = 0$ if and only if $x \perp y$, and hence for any $M \in \mathscr{L}$, $y \in M^{\perp}$ if and only if $y \in \xi(M)$. This proves that $\perp$ and ξ coincide. The proof of theorem 4.1 is complete.

Theorem 4.1 shows that a logic may be asserted to be standard as soon as it is identified as a member of an apparently more general class of lattices. We have proved theorem 4.1 to emphasize our point of view that the important role played by standard logics is not due to accident but stems from deeper mathematical reasons, and that theorems such as the present one lead to a clarification of some of these reasons. This said, the stage is set for the detailed analysis of the standard logics and we proceed to do it now.

2. THE STANDARD LOGICS: OBSERVABLES AND STATES

We now proceed to describe the structure of the set of observables and the set of states associated with a standard logic and examine the manner in which the general concepts of Chapter III specialize. Throughout this section we write $\mathbf{D}$ to denote one of $\mathbf{R}$, $\mathbf{C}$, or $\mathbf{H}$. Let $\mathscr{H}$ be a separable infinite dimensional Hilbert space over $\mathbf{D}$. We write $\mathscr{L}$ for the logic $\mathscr{L}(\mathscr{H},\mathbf{D})$ of all closed linear manifolds of $\mathscr{H}$ (with its canonical orthocomplementation).

For any closed linear manifold M of $\mathscr{H}$, let us write P^M for the orthogonal projection on M. If $f(E \to f(E))$ is any observable associated with the logic $\mathscr{L}$, then the map $E \to P^{f(E)}$ is a *projection valued measure*† based on $\mathscr{B}(\mathbf{R}^1)$. Conversely, it is obvious that to any projection valued measure $P(E \to P_E)$ based on $\mathscr{B}(\mathbf{R}^1)$ there exists a unique observable $f(E \to f(E))$ such that $P_E = P^{f(E)}$ for all real Borel sets E.

† The concept of a projection valued measure is nothing new; it is a standard tool of spectral theory. For complex Hilbert spaces we refer the reader to sections 35–38 of Halmos [2]. For the real case very few changes have to be made. For the quaternionic case and for projection valued measures defined *on the real line*, the theory is essentially the same, since the real field is the center of $\mathbf{H}$. We shall assume the basic facts concerning projection valued measures as known.

Lemma 4.9. *Let* $M_1, M_2 \in \mathscr{L}$. *Then* $M_1 \leftrightarrow M_2$ *if and only if the projections* P^{M_1} *and* P^{M_2} *commute.*

Proof. Let $M_1, M_2 \in \mathscr{L}$. Assume first that $M_1 \leftrightarrow M_2$. Then there exists mutually orthogonal elements N_1, N_2, N of $\mathscr{L}$ such that $M_i = N_i + N$ $(i=1, 2)$. Clearly, $P^{M_i} = P^{N_i} + P^N$, and one is led immediately to the equation $P^{M_1}P^{M_2} = P^{M_2}P^{M_1}$. Conversely, let P^{M_1} and P^{M_2} commute and let $P = P^{M_1}P^{M_2}$. Then P is a projection. If we write $Q_i = P^{M_i} - P$, then it is easily verified that Q_i is a projection $(i=1, 2)$ and $PQ_i = Q_iP = 0$, $Q_1Q_2 = Q_2Q_1 = 0$. If N_1, N_2, and N are the closed linear manifolds which are the ranges of Q_1, Q_2, and P, respectively, then it follows easily that N_1, N_2, and N are mutually orthogonal and $M_i = N_i \vee N$ $(i=1, 2)$. This proves that $M_1 \leftrightarrow M_2$.

Lemma 4.10. *Let* X *be a set and* $\mathscr{B}$ *a* σ*-algebra of subsets of* X. *Let* $P(E \to P_E)$ *be a projection valued measure in* $\mathscr{H}$ *based on* $\mathscr{B}$ *and* $\mathscr{F}$ *the algebra (over* **R**) *of all real valued bounded measurable functions on* X. *For* $g \in \mathscr{F}$ *let* $\|g\| = \sup\{|g(x)| : x \in X\}$. *Then, for any* $g \in \mathscr{F}$, *there exists a unique, bounded, self-adjoint operator* B_g *on* $\mathscr{H}$ *such that*

$$\langle B_g u,v\rangle = \int_X g(x) d\nu_{u,v}(x) \qquad (u, v \in \mathscr{H}), \tag{22}$$

$$\left(B_g = \int_X g dP, \qquad \textit{in symbols}\right)$$

where $\nu_{u,v}$ *is the* **D**-*valued measure on* $\mathscr{B}$ *defined by*

$$\nu_{u,v}(E) = \langle P_E u,v\rangle.$$

The spectral measure of B_g *is the projection valued measure* $E \to P_{g^{-1}(E)}$. *One has:*

$$\|B_g\| \leq \|g\| \qquad (g \in \mathscr{F}). \tag{23}$$

The correspondence $g \to B_g$ *is a homomorphism from the algebra* $\mathscr{F}$ *into the algebra (over* **R**) *of all bounded operators on* $\mathscr{H}$. *In particular,* B_{g_1} *and* B_{g_2} *commute for all* $g_1, g_2 \in \mathscr{F}$.

Proof. The proofs are similar to the ones in section 37 of Halmos [2] and we omit them. In the quaternionic case we have to use the fact that the real field is the center of **H**.

Theorem 4.11. *Let* **D** *be one of* **R**, **C**, *or* **H** *and* $\mathscr{L}$ *the logic of all closed linear manifolds of* $\mathscr{H}$. *For any observable* $f(E \to f(E))$ *associated with*

$\mathscr{L}$, let A_f be the self-adjoint (not necessarily bounded) linear operator with spectral measure $E \to P^{f(E)}$. Then

$$f \to A_f \tag{24}$$

is a one-one correspondence between the set of all observables associated with $\mathscr{L}$ and the set of all self-adjoint operators on $\mathscr{H}$. The observable f is bounded if and only if the operator A_f is bounded. Two bounded observables f_1 and f_2 are simultaneously observable if and only if the (bounded) operators A_{f_1} and A_{f_2} commute. If f is any bounded observable and p is a polynomial with real coefficients (in one variable), then the operator corresponding under (24) *to $p \circ f$ is $p(A_f)$. More generally, if $f_1, \cdots, f_r$ are bounded observables, any two of which are simultaneously observable, and if p is a real polynomial of r real variables, then the observable $p \circ (f_1, \ldots, f_r)$ (cf. theorem* 3.17) *has the corresponding operator $p(A_{f_1}, \ldots, A_{f_r})$.*

Proof. The first assertion concerning the correspondence $f \to A_f$ is an immediate consequence of the spectral theorem. Next, an observable f is bounded if and only if for some compact set $K \subseteq \mathbf{R}^1$, $f(K) = \mathscr{H}$. This is possible if and only if $P^{f(K)} = I$, which in turn is well known to be equivalent to the requirement that A_f be bounded. Let now f_1 and f_2 be bounded observables. Then f_1 and f_2 are simultaneously observable if and only if $f_1(E_1) \leftrightarrow f_2(E_2)$ for all pairs of Borel sets E_1 and E_2. From lemma 4.9 and standard spectral theory, it now follows that $f_1 \leftrightarrow f_2$ if and only if A_{f_1} and A_{f_2} commute. We now come to the last pair of assertions. Since the first of these is a special case of the second for $r = 1$, we shall confine ourselves to the proof of the second. Let $f_1, \cdots, f_r$ be bounded observables such that $f_i \leftrightarrow f_j$ for all $i, j = 1, 2, \cdots, r$. By theorem 3.17 there exists a σ-homomorphism $F \to f(F)$ of $\mathscr{B}(R^r)$ into $\mathscr{L}$ such that for any i $(1 \leq i \leq r)$ and any $E \in \mathscr{B}(\mathbf{R}^1)$,

$$f_i(E) = f(\pi_i^{-1}(E)),$$

π_i being the projection $(t_1, \cdots, t_r) \to t_i$ of $\mathbf{R}^r$ onto $\mathbf{R}^1$. Let $P_F = P^{f(F)}$. Then $F \to P_F$ is a projection valued measure on $\mathscr{B}(\mathbf{R}^r)$. Since the f_i are bounded observables, there exists a compact set $K \subseteq \mathbf{R}^1$ such that $f_i(K) = \mathscr{H}$ for $i = 1, 2, \cdots, r$. We now apply lemma 4.10 by taking $X = K^r$, $\mathscr{B} = \mathscr{B}(K^r)$. For any $E \in \mathscr{B}(K)$, $P_{\pi_i^{-1}(E)} = P^{f_i(E)}$, and so, writing

$$\nu_{u,v}(F) = \langle P_F u,v \rangle, \qquad \nu^i_{u,v}(E) = \langle P^{f_i(E)} u,v \rangle,$$

we get

$$\begin{aligned} \langle B_{\pi_i} u,v \rangle &= \int_{K^r} \pi_i(t_1, t_2, \cdots, t_r) d\nu_{u,v}(t_1, \cdots, t_r) \\ &= \int_K t d\nu^i_{u,v}(t) \\ &= \langle A_{f_i} u,v \rangle. \end{aligned}$$

Thus, $B_{\pi_i} = A_{f_i}$, $i = 1, 2, \cdots, r$. Since $g \to B_g$ is a homomorphism, we have, for any real polynomial p of $t_1, \cdots, t_r$,

$$B_p = p(B_{\pi_1}, \cdots, B_{\pi_r}) = p(A_{f_1}, \cdots, A_{f_r}).$$

But the self-adjoint operator B_p has the spectral measure $E \to P_{p^{-1}(E)}$ by lemma 4.10. B_p therefore coincides with the operator corresponding to the observable $E \to f(p^{-1}(E))$ which is precisely $p \circ (f_1, \cdots, f_r)$. This proves the theorem completely.

Remark. That $f_1 \leftrightarrow f_2$ is equivalent to the equation $A_{f_1}A_{f_2} = A_{f_2}A_{f_1}$ is a celebrated result of von Neumann. The fact that $p \circ (f_1, \cdots, f_r)$ corresponds to the operator $p(A_{f_1}, \cdots, A_{f_r})$ is remarkable since it links the *algebra of functions* of $f_1, \cdots, f_r$ with the *algebra of operators* on $\mathscr{H}$.

We now take up the problem of determining the states of the logic $\mathscr{L}$. Our basic result is the theorem of Gleason [1] which asserts that every state of $\mathscr{L}$ can be described in a canonical way by what is usually known as a *density matrix*. We shall now proceed to a detailed discussion of this theorem.

To any vector $u \in \mathscr{H}$ with $\|u\| = 1$, we associate the mapping p_u of $\mathscr{L}$ into the set of nonnegative real numbers by defining

$$p_u(M) = \langle P^M u, u \rangle = \|P^M u\|^2 \qquad (M \in \mathscr{L}). \tag{25}$$

If $M = \mathscr{H}$, $p_u(M) = \|u\|^2 = 1$, while $p_u(0) = 0$. If $M_1, M_2, \cdots$ is a sequence of elements of $\mathscr{L}$ such that $M_i \perp M_j$ $(i \neq j)$ then, writing $M = \bigvee_i M_i$, we have:

$$\begin{aligned} p_u(M) &= \|P^M u\|^2 \\ &= \sum_i \|P^{M_i} u\|^2 \\ &= \sum_i p_u(M_i). \end{aligned}$$

In other words, p_u is a state of $\mathscr{L}$. It is clear that if $c \in \mathbf{D}$ and $|c| = 1$, $p_{cu} = p_u$.

The general description of states depends on the concept of an operator of trace class. On complex Hilbert spaces they are defined and treated in many books on Hilbert space theory (cf. Dunford and Schwartz [2], for example). The treatment in the real and quaternionic cases is similar. A bounded operator A of $\mathscr{H}$ is of *trace class* if, for any orthonormal basis $\{e_i\}$, the series

$$\sum_i |\langle Ae_i, e_i \rangle| < \infty;$$

then the sum

$$\operatorname{tr}(A) = \sum_i \langle Ae_i, e_i \rangle$$

exists for any orthonormal basis $\{e_i\}$ and is independent of the basis used; it is called the trace of A. If A is of trace class and B is *any* bounded operator, AB and BA are of trace class; and moreover,

$$\operatorname{tr}(AB) = \operatorname{tr}(BA).$$

If A and B are of trace class, so are $\alpha A + \beta B$ where α, β are in the center of the division ring over which $\mathscr{H}$ is defined. We shall assume the reader to be familiar with the basic properties of these operators.

The states of the form $p_u (u \in \mathscr{H}, \|u\| = 1)$ can also be described in another way. Let U be the projection on the one-dimensional linear manifold $\mathbf{D} \cdot u$. Then, for any bounded operator A, AU is a bounded operator, which is of trace class, since U is (trivially) of trace class. Let $\{e_n\}$ be an orthonormal basis for $\mathscr{H}$ with $u = e_1$. Then

$$\begin{aligned} \operatorname{tr}(AU) &= \operatorname{tr}(UA) \\ &= \sum_i \langle AUe_i, e_i \rangle \\ &= \langle Au, u \rangle. \end{aligned} \tag{26}$$

In particular,

$$p_u(M) = \operatorname{tr}(P^M U). \tag{27}$$

Suppose that $u_1, u_2, \cdots$ is any sequence of vectors in $\mathscr{H}$ with $\|u_i\| = 1$ for all i. Let $a_1, a_2, \cdots$ be a sequence of nonnegative numbers with $a_1 + a_2 + \cdots = 1$. Let U_i be the projection on the one-dimensional manifold $\mathbf{D} \cdot u_i$. Then

$$U = \sum_i a_i U_i \tag{28}$$

is a well-defined, bounded, nonnegative, self-adjoint operator of trace class and

$$\operatorname{tr}(U) = 1. \tag{29}$$

If $M \in \mathscr{L}$, we have

$$\begin{aligned} \operatorname{tr}(P^M U) &= \sum_i a_i \operatorname{tr}(P^M U_i) \\ &= \sum_i a_i p_{u_i}(M). \end{aligned}$$

Thus the function p_U defined by

$$p_U(M) = \operatorname{tr}(P^M U) \tag{30}$$

is a state of $\mathscr{L}$ and in fact is the state $\sum_i c_i p_{u_i}$. Conversely, let U be a bounded, self-adjoint operator such that (i) U is nonnegative, i.e., $\langle Ux, x \rangle \geq 0$ for all $x \in \mathscr{H}$, and (ii) U is of trace class and $\operatorname{tr}(U) = 1$. Then, the function p_U defined by (30) is a state of $\mathscr{L}$. We call the operators U which are bounded, self-adjoint, nonnegative, and of trace class, *von Neumann operators*.

The question naturally arises whether every state of $\mathscr{L}$ is of the form p_U for a suitable von Neumann operator U of trace 1. It was Gleason [1] who first proved that this is indeed so. His result, which is a cornerstone of the mathematical foundations of quantum mechanics, is undoubtedly one of the most profound theorems in this subject.

Gleason's theorem is quite complicated to prove and we need to develop a number of technical lemmas. We begin with the concept of a frame function. Let $\mathscr{E}$ be the unit sphere of $\mathscr{H}$, i.e.,

$$\mathscr{E} = \{x : x \in \mathscr{H}, \|x\| = 1\}. \tag{31}$$

A real valued function $f(x \rightarrow f(x))$ defined on $\mathscr{E}$ is a *frame function* if

(32)
(i) $f(cx) = f(x)$ for $x \in \mathscr{E}$ and $c \in \mathbf{D}$ with $|c| = 1$,
(ii) there exists a constant W such that $\sum_n f(e_n) = W$ for any orthonormal basis $\{e_n\}$ of $\mathscr{H}$ (recall that $\mathscr{H}$ is separable).

The constant W is called the *weight* of f.

Suppose that U is a bounded, self-adjoint operator of trace class. Let us define the function f_U by

$$f_U(x) = \langle Ux,x\rangle \qquad (x \in \mathscr{E}). \tag{33}$$

Then f_U is obviously a frame function whose weight is tr U. A frame function f is called *regular* if there exists a bounded, self-adjoint operator U such that

$$f = f_U. \tag{34}$$

Clearly, if f is regular, there exists only one bounded, self-adjoint operator U such that (34) holds; for, if U and V satisfy (34), $\langle (U-V)x, x\rangle = 0$ for all $x \in \mathscr{H}$ and hence $U - V = 0$.

Suppose that μ is a state of $\mathscr{L}$. For any vector $x \in \mathscr{E}$ let us define

$$f_\mu(x) = \mu(\mathbf{D}\cdot x). \tag{35}$$

Then f_μ is a *nonnegative* frame function of weight 1. If U is a von Neumann operator of trace 1 and $\mu = p_U$ the state defined by (30), a simple calculation shows

$$f_\mu = f_U, \tag{36}$$

i.e., f_μ is regular. Conversely, if μ is a state of $\mathscr{L}$ and if we assume that

$f_\mu = f_U$ for a bounded, self-adjoint operator, then it can be shown that U is a von Neumann operator of trace 1, and $\mu = p_U$. In fact, as

$$\langle Ux,x\rangle = f_\mu(x) \geq 0,$$

U is nonnegative; then, if $\{e_n\}$ is any orthonormal basis of $\mathscr{H}$,

$$\sum_n \langle Ue_n,e_n\rangle = \sum_n f_\mu(e_n) = 1$$

so that U is of trace class and has unit trace. This shows that the states μ and p_U coincide for all one-dimensional subspaces of $\mathscr{H}$ and hence coincide for all elements of $\mathscr{L}$. In other words, Gleason's theorem is entirely equivalent to proving that every nonnegative frame function is regular.

We shall first examine the special case when $\mathbf{D}=\mathbf{R}$ and when $\mathscr{H}$ is finite dimensional. If $\mathscr{H}$ is the real line, then the only frame functions are constants. If $\mathscr{H}$ is the plane $\mathbf{R}^2$, then it is obvious that a frame function can be arbitrarily defined on any quadrant of the unit circle and hence need have no special properties.

Lemma 4.12. *Let* $\mathscr{H} = \mathbf{R}^2$, *the plane* ($\mathbf{D}$ = *the real field*) *and let* $\mathscr{E}$ *be described as*

$$\mathscr{E} = \{(\cos\theta, \sin\theta) : 0 \leq \theta < 2\pi\}. \tag{37}$$

Then the function $f : (\cos\theta, \sin\theta) \to \cos n\theta$, *$n$ being an integer, is a frame function if and only if either* $n=0$ *or* $n \equiv 2$ (*mod* 4).

Proof. In order that f be a frame function we must have, for some constant W,

$$\cos n\theta + \cos\left(n\left(\theta+\frac{\pi}{2}\right)\right) = W \tag{38}$$

for all θ. Differentiating with respect to θ, we have:

$$n\left\{\sin n\theta + \sin n\left(\theta+\frac{\pi}{2}\right)\right\} = 0$$

for all θ, i.e.,

$$n\left\{(\sin n\theta)\left(1+\cos\frac{n\pi}{2}\right)+\sin\frac{n\pi}{2}\cos n\theta\right\} = 0.$$

This is possible if and only if either $n=0$ or $\cos(n\pi/2) = -1$, i.e., if and only if $n=0$ or $n\equiv 2$ modulo 4.

Lemma 4.13. *Let* $\mathscr{H} = \mathbf{R}^3$ *be the three dimensional real Euclidean space. Then every continuous frame function on the unit sphere* $\mathscr{E}$ *is regular.*

Proof. The proof makes use of harmonic analysis on $\mathscr{E}$ (cf. Weyl [1], pp. 60–63, 142–146). Let $\mathscr{C}$ be the set of all real valued continuous functions on $\mathscr{E}$ and G the group of all orthogonal transformations of R^3 which have determinant $+1$. For any element h of $\mathscr{C}$ and any $a \in G$, we define h^a by

$$h^a(x) = h(a^{-1}x) \qquad (x \in \mathscr{E}). \tag{39}$$

A linear manifold $\mathscr{M} \subseteq \mathscr{C}$ is said to be *invariant* if $h^a \in \mathscr{M}$ for all $a \in G$ whenever $h \in \mathscr{M}$. Then, according to classical spherical harmonic analysis, one has the following description of $\mathscr{C}$: (1) for each integer $n = 0, 1, \cdots$ there is a $2n+1$ dimensional subspace $\mathscr{F}_n$ of $\mathscr{C}$ which is invariant and no proper linear submanifold of which is invariant, i.e., $\mathscr{F}_n$ is irreducible under the action (39) of the group G; (2) an element $h \in \mathscr{C}$ belongs to $\mathscr{F}_n$ if and only if it is the restriction to $\mathscr{E}$ of a polynomial p of x, y, z (the coordinates on R^3) with the properties

$$\text{(40)} \quad \begin{array}{ll} \text{(i)} & p \text{ is homogeneous in } x, y, z \text{ of degree } n, \\ \text{(ii)} & \left(\dfrac{\partial^2}{\partial x^2}+\dfrac{\partial^2}{\partial y^2}+\dfrac{\partial^2}{\partial z^2}\right)p = 0; \end{array}$$

(3) every continuous function on $\mathscr{E}$ is the uniform limit of a sequence of functions each of which is a finite linear combination of elements of the $\mathscr{F}_n$; and finally (4) if $\mathscr{F}$ is any invariant linear manifold, closed under the uniform convergence topology of $\mathscr{C}$, and $n_1, n_2, \cdots$ $(0 \leq n_1 < n_2 < \cdots)$ are all the integers such that $\mathscr{F}_{n_j} \subseteq \mathscr{F}$, $j = 1, 2, \cdots$, then

$$\mathscr{F} = \Big(\sum_j \mathscr{F}_{n_j}\Big)^{-}, \tag{41}$$

where $\sum$ denotes the algebraic sum, and the bar denotes the closure.

Let now $\mathscr{F}$ be the set of all continuous frame functions on $\mathscr{E}$. Evidently G-transforms (cf. (39)) and uniform limits of continuous frame functions are continuous frame functions. Thus the set $\mathscr{F}$ is a closed linear manifold in $\mathscr{C}$ which is invariant. Therefore, there exist integers $n_1, n_2, \cdots$ with the properties described in (41). We now claim that if n is any integer other than 0 and 2, then, $\mathscr{F}_n \nsubseteq \mathscr{F}$. In fact, let n be any integer. We use cylindrical coordinates ρ, θ, z on $\mathscr{E}$ defined by

$$x = \rho \cos \theta, \quad y = \rho \sin \theta, \quad z = z.$$

Then it is well known that

$$\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2}+\frac{\partial^2}{\partial z^2} = \frac{\partial^2}{\partial \rho^2}+\frac{1}{\rho^2}\frac{\partial^2}{\partial \theta^2}+\frac{1}{\rho}\frac{\partial}{\partial \rho}+\frac{\partial^2}{\partial z^2}.$$

One can easily check that the functions

$$\rho^n \cos n\theta,\ [\rho^n - 2(n-1)z^2\rho^{n-2}] \cos (n-2)\theta$$

are homogeneous polynomials of degree n which satisfy the equation

$$\left(\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2}+\frac{\partial^2}{\partial z^2}\right)p = 0.$$

Thus if $\mathscr{F}_n \subseteq \mathscr{F}$, then $\rho^n \cos n\theta$ and $[\rho^n - 2(n-1)z^2\rho^{n-2}]\cdot\cos(n-2)\,\theta$ are frame functions. For any θ, $(\cos\theta, \sin\theta, 0)$, $(-\sin\theta, \cos\theta, 0)$ and $(0,0,1)$ form a frame of R^3 so that the functions obtained by putting $z=0$ are frame functions on the unit circle of the x, y plane. Thus $\cos n\theta$ and $\cos(n-2)\theta$ are both frame functions. By lemma 4.12 this is not possible unless $n=0$ or 2. Thus we must have $\mathscr{F} \subseteq (\mathscr{F}_0 + \mathscr{F}_2)^-$ i.e., $\mathscr{F} \subseteq \mathscr{F}_0 + \mathscr{F}_2$, since $\mathscr{F}_0$ and $\mathscr{F}_2$ are finite dimensional. We now claim that $\mathscr{F} = \mathscr{F}_0 + \mathscr{F}_2$ and that every element of $\mathscr{F}_0 + \mathscr{F}_2$ is a regular frame function. In fact, $\mathscr{F}_0$ consists of the constants which are regular. On the other hand, $\mathscr{F}_2$ consists of the quadratic forms $q(x,y,z)$ which satisfy

$$\left(\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2}+\frac{\partial^2}{\partial z^2}\right)q = 0,$$

i.e.,

$$q(x,y,z) = ax^2+by^2+cz^2+2fyz+2gzx+2hxy$$

with

$$a+b+c = 0.$$

Thus a function of $\mathscr{C}$ lies in $\mathscr{F}_0 + \mathscr{F}_2$ if and only if it is of the form

$$x \to \langle Ax,x\rangle$$

(x is a vector in $\mathscr{E}$), where A is a symmetric matrix of order 3. In other words, every element of $\mathscr{F}_0 + \mathscr{F}_2$ is a regular frame function. Thus $\mathscr{F} = \mathscr{F}_0 + \mathscr{F}_2$. This completes the proof that every continuous frame function is regular.

We now take up the proof of the deep lemma that every nonnegative frame function on the unit sphere in real three-dimensional space is actually continuous. Gleason's proof of this result is an intricate one based on spherical geometry. We shall first introduce some terminology. We utilize the usual latitude-longitude coordinates on a sphere. For any point p of the unit sphere $\mathscr{E}$ of R^3, we define the *northern hemisphere relative to* p to be all the points whose latitude θ relative to p as north pole satisfies $0 < \theta < \pi/2$. We shall denote this by N_p. Given p, let q be some point of N_p. Clearly, $q \neq p$. Among the great circles which pass through this point q there is one whose plane cuts the equator relative to p at points which are orthogonal to q. If the latitude of q is θ, it can be shown that this great circle passes through only those points whose latitudes φ satisfy $-\theta \leq \varphi \leq \theta$. It is therefore appropriate to call this the *EW great circle* through q.

The essence of our argument is contained in the next lemma. To motivate this lemma, let us consider in the usual coordinates x, y, z the function

$$(x, y, z) \rightarrow x^2 + y^2.$$

This is a frame function which is nonnegative and vanishes at the north pole $p = (0,0,1)$. It is constant on points with the same latitude. For any point q of the northern hemisphere, the points of N_p other than q on the EW great circle through q, have latitudes strictly less than the latitude of q, and consequently this frame function reaches its minimum on the EW great circle through q, at q. The next lemma asserts that every frame function which is nonnegative has this property *approximately*.

Lemma 4.14. *Let $p \in \mathscr{E}$ and g be a nonnegative frame function which is constant on the equator opposite to p. Then for any q in N_p we have:*

$$g(q) \leq g(x) + g(p) \tag{42}$$

for every point x on C, the EW great circle through q.

Proof. We first assert that if r is any point of N_p, then

$$g(r) \leq k + g(p), \tag{43}$$

where k is the constant value which g takes on the equator. To see this, consider the EW great circle through r. If r' is one of the points at which this meets the equator, then $r \perp r'$ and hence r, r' are two points of a possible frame. The frame function g being nonnegative, one has:

$$g(r) + g(r') \leq W,$$

where W is the weight of g. But if we take any two points a, a' on the equator which are orthogonal, a, a', p is a frame and, consequently,

$$2k + g(p) = W.$$

Thus

$$g(r) \leq W - g(r') \leq 2k + g(p) - k \leq k + g(p).$$

Suppose now q is an arbitrary point in the northern hemisphere. The inequality just now established proves the lemma if x lies on the equator. For x below the equator, $-x$ lies above the equator, and, since $g(u) = g(-u)$ for all u, it is enough to prove the inequality (42) when x lies on the EW great circle C through q and is also situated in the northern hemisphere.

Let y be a point on C, orthogonal to x, and lying in the northern hemisphere. We then have, if q' is a point where C meets the equator,

$$g(x)+g(y) = g(q)+g(q') = g(q)+k$$

and, since $g(y) \leq k+g(p)$ by (43), one has:

$$\begin{aligned} g(x) &= g(q)+k-g(y) \\ &\geq g(q)-g(p), \end{aligned}$$

which establishes the lemma.

Lemma 4.15. *With p and g as in lemma 4.14, let z be a fixed point in N_p. Let M_z be the set of points x on N_p which have the property that for some point y of N_p the EW great circle through x contains y and the EW great circle through y contains z. Then the interior of M_z is nonempty. Moreover, for any x of M_z, one has:*

$$g(x) \leq g(z)+2g(p). \tag{44}$$

Proof. Since $g(x) \leq g(y)+g(p)$ and $g(y) \leq g(z)+g(p)$, the required inequality follows for all the points x in question. It now remains to establish that M_z contains a nonempty open subset of N_p. We use Cartesian coordinates. We can arrange matters so that the coordinates of p are $(0,0,1)$, while the coordinates of z are $(\cos\alpha, 0, \sin\alpha)$. We have $0<\alpha<\pi$, $\alpha \neq \pi/2$. One can easily see that for any point y whose coordinates are (ξ,η,ζ) $(0<\zeta<1)$ the EW great circle through y contains z if and only if

$$(\xi^2+\eta^2)\sin\alpha - \xi\zeta\cos\alpha = 0.$$

Consider now the function ψ, where

$$\psi(\xi,\eta,\zeta) = (\xi^2+\eta^2)\sin\alpha - \xi\zeta\cos\alpha.$$

This is greater than 0 on the equator. On the other hand, there are points of N_p for which it is <0. Therefore, if we define A to be the set of points

$$A = \{(\xi,\eta,\zeta) : 0<\zeta<1, \quad \psi(\xi,\eta,\zeta) < 0\},$$

then A is a nonempty open subset of N_p. Suppose x lies in A. Then $\psi(x)<0$ while the EW great circle through x meets the equator at a point at which ψ is positive. Since ψ is continuous, there must be a point y on the EW great circle through x at which ψ vanishes; thus $x \in M_z$. In other words, $A \subseteq M_z$. This proves the lemma.

Lemma 4.16. *With p and g as in lemma 4.14, let $\eta>0$ be a positive number such that $g(p)<\eta$. Then there exists a point q of N_p and an open set M around q such that*

$$0 \leq \sup_{u\in M} g(u) - \inf_{u\in M} g(u) \leq 3\eta. \tag{45}$$

Proof. Define $b = \inf_{u \in N_p} g(u)$. Then $b \geq 0$. Choose $z \in N_p$ such that $b \leq g(z) \leq b + \eta$. Consider the set M_z. By lemma 4.15, M_z has a nonempty interior and for every point x of M_z one has:

$$g(x) \leq g(z) + 2\eta,$$

that is,

$$b \leq g(x) \leq b + 3\eta.$$

Let q be any point in the interior M of M_z. Then M is an open neighborhood of q and

$$0 \leq \sup_{u \in M} g(u) - \inf_{u \in M} g(u) \leq 3\eta.$$

This proves the lemma.

Lemma 4.17. *Let f be any nonnegative frame function and r any point such that for some open neighborhood M of r we have:*

$$0 \leq \sup_{u \in M} f(u) - \inf_{u \in M} f(u) \leq \alpha. \tag{46}$$

Then for every point s of $\mathscr{E}$ there is an open neighborhood M_s of s such that

$$0 \leq \sup_{u \in M_s} f(u) - \inf_{u \in M_s} f(u) \leq 4\alpha. \tag{47}$$

Proof. Use coordinates so that r appears as the north pole. Let t be any point on the equator opposite r. Since M is an open neighborhood of r, we can assume that every point whose latitude is $\geq (\pi/2) - \theta_0$ belongs to M, θ_0 being a suitable angle in $]0, \pi/2[$. Consider now a point u due south of t and latitude $-\frac{1}{2}\theta_0$. Then by continuity it is clear that there exists an open set M_t around t with the following property: if k is any point of M_t, then, on the great circle through u and k, there are points orthogonal to u and k, respectively, which are situated in M. Let k, k' be two points of M_t, and C, C' the respective great circles through u, k and u, k'. Let x, y and x', y' be pairs of points, respectively, on these two great circles such that

$$x, y, x', y' \in M,$$
$$x \perp u, \quad y \perp k, \quad x' \perp u, \quad y' \perp k'.$$

We then have:

$$f(x) + f(u) = f(k) + f(y),$$
$$f(x') + f(u) = f(k') + f(y').$$

Subtracting and rearranging, we have:

$$\begin{aligned} |f(k) - f(k')| &\leq |f(x) - f(x')| + |f(y) - f(y')| \\ &\leq 2\alpha, \end{aligned}$$

so that

$$0 \leq \sup_{u \in M_t} f(u) - \inf_{u \in M_t} f(u) \leq 2\alpha.$$

If s is any point of $\mathscr{E}$, there exists a point t of $\mathscr{E}$ such that t is on the equator opposite r and s is on the equator opposite t. Applying the preceding argument twice, we see that there is an open set M_s around s such that

$$0 \leq \sup_{u\in M_s} f(u) - \inf_{u\in M_s} f(u) \leq 4\alpha.$$

This proves the lemma.

Lemma 4.18. *Let f be a nonnegative frame function on $\mathscr{E}$. Then f is continuous.*

Proof. Let $\varepsilon > 0$ be a given positive number. We shall prove that every point of $\mathscr{E}$ has a neighborhood in which the oscillation of f is at most ε. Since a constant is a continuous frame function, we can, by subtracting the infimum of f (over $\mathscr{E}$) from f, assume without any loss of generality that

$$\inf_{u\in\mathscr{E}} f(u) = 0.$$

Let $\eta > 0$ and p a point such that $f(p) < \eta/2$. Let σ be the rotation around the z-axis with p as north pole through an angle $\pi/2$. Let the frame function g be defined by

$$g(x) = f(x) + f(\sigma x).$$

g is nonnegative, and at p we have:

$$g(p) < \eta;$$

moreover, g is constant on the equator. By lemma 4.16 there exists a point q of the northern hemisphere and some neighborhood M_q of q such that

$$0 \leq \sup_{u\in M_q} g(u) - \inf_{u\in M_q} g(u) \leq 3\eta.$$

By lemma 4.17 there exists an open set M_p around p such that

$$0 \leq \sup_{u\in M_p} g(u) - \inf_{u\in M_p} g(u) \leq 12\eta.$$

Since $g(p) < \eta$, we have:

$$0 \leq \sup_{u\in M_p} g(u) \leq 13\eta.$$

Since $0 \leq f(u) \leq g(u)$ for all u, we have:

$$0 \leq \sup_{u\in M_p} f(u) \leq 13\eta$$

and hence we have:

$$0 \leq \sup_{u\in M_p} f(u) - \inf_{u\in M_p} f(u) \leq 13\eta.$$

Applying lemma 4.17 again, we see that every point of $\mathscr{E}$ has a neighborhood over which the oscillation of f is at most 52η. If we choose $\eta = \varepsilon/52$, we are through.

Lemma 4.19. *Let $\mathscr{H}$ be a real separable Hilbert space of dimension at least 3. Then every nonnegative frame function f on the unit sphere $\mathscr{E}$ of $\mathscr{H}$ is regular.*

Proof. Since every two-dimensional subspace L of $\mathscr{H}$ can be imbedded in a three-dimensional subspace, it follows from lemma 4.18 that the restriction of f to the unit sphere $\mathscr{E}(L)$ of L is regular, and consequently there is a *unique symmetric* bilinear form on $L \times L$, say B_L, such that

$$B_L(u,u) = f(u) \tag{48}$$

for all $u \in \mathscr{E}(L)$. We shall now define a real valued function B on $\mathscr{H} \times \mathscr{H}$ as follows. Let x and y be two vectors of $\mathscr{H}$ and L a two-dimensional subspace containing x and y. We define

$$B(x,y) = B_L(x,y). \tag{49}$$

Clearly, $B(x,y)$ is well defined if x and y are linearly independent, since in this case L is unique; or if one or both of them is zero, since we then have $B_L(x,y) = 0$ for any such L. If x and y were dependent and neither is zero, then they span a one-dimensional subspace. If L_1 and L_2 are two two-dimensional subspaces containing x and y, then both L_1 and L_2 are contained in a three-dimensional subspace, say M. Let C be a symmetric bilinear form on $M \times M$ such that

$$C(a,a) = f(a)$$

for all vectors a in $\mathscr{E}(M)$. The restrictions of C to $L_1 \times L_1$ and $L_2 \times L_2$ are symmetric bilinear forms whose quadratic forms coincide with f on the unit spheres of L_1 and L_2, respectively. Since B_{L_1} and B_{L_2} also have this property, we conclude that

$$B_{L_1}(x,y) = C(x,y),$$
$$B_{L_2}(x,y) = C(x,y),$$

which proves that $B(x,y)$ is well defined. Moreover, we have

$$B(x,x) = f(x) \tag{50}$$

for all unit vectors x in $\mathscr{H}$. We now claim that B is bilinear. The homogeneity of B as well as its symmetry involves only two vectors and so follows from the homogeneity and symmetry of B_L. To complete the proof of the bilinearity (in view of the symmetry of B), it is enough to show that

$$B(x, y+z) = B(x,y) + B(x,z)$$

for any three vectors x, y, z. Let M be a three-dimensional subspace containing x, y and z. Let L_1, L_2, L_3 be two-dimensional subspaces of M containing, respectively, x, y; x, z; and x, $y+z$. Let C be a symmetric bilinear form on $M \times M$ such that $C(u,u)=f(u)$ for all unit vectors u in M. Then we see at once that

$$C(x,y) = B_{L_1}(x,y), \qquad C(x,z) = B_{L_2}(x,y)$$

and

$$C(x, y+z) = B_{L_3}(x, y+z).$$

Since C is bilinear, we have:

$$\begin{aligned} B(x, y+z) &= B_{L_3}(x, y+z) \\ &= C(x, y+z) \\ &= C(x,y)+C(x,z) \\ &= B_{L_1}(x,y)+B_{L_2}(x,z) \\ &= B(x,y)+B(x,z), \end{aligned}$$

which proves the bilinearity of B.

Since

$$0 \leq B(x,x) \leq 1$$

for all x with $\|x\|=1$, $x, y \to B(x,y)$ is a *bounded* bilinear form. Thus there exists a bounded, self-adjoint operator U such that for all x, y in $\mathscr{H}$

$$\begin{aligned} B(x,y) &= \langle Ux,y\rangle, \\ f(x) &= \langle Ux,x\rangle. \end{aligned}$$

Thus we have proved that f is regular.

We now proceed to the case of Hilbert spaces over **D**, where **D** may either be the complex number field or the quaternion division ring. Let $\mathscr{H}$ be a Hilbert space over **D**. A closed **R**-linear manifold S of $\mathscr{H}$ is called *completely real* if the inner product $\langle . , . \rangle$ in $\mathscr{H} \times \mathscr{H}$ takes real values on $S \times S$; equivalently, if and only if there is an orthonormal set $\{e_j\}$ such that S is the closure of the real linear combinations of the e_j.

Lemma 4.20. *Let $\mathscr{R}$ be a real separable Hilbert space, and f_0 a regular nonnegative frame function of weight W. Then for all $a, b \in \mathscr{R}$ with $\|a\|=\|b\|=1$,*

$$|f_0(a)-f_0(b)| \leq 2W\|a-b\|. \tag{51}$$

Proof. There is an operator T, of trace class and self-adjoint such that $\langle Tc,c\rangle=f_0(c)$ for all unit vectors c. Then

$$\begin{aligned} |f_0(a)-f_0(b)| &= |\langle Ta,a\rangle-\langle Tb,b\rangle| \\ &= |\langle T(a-b), (a+b)\rangle| \\ &\leq 2\|T\|\cdot\|a-b\|. \end{aligned}$$

But $0 \leq \langle Tc, c \rangle \leq W$ for all unit vectors c so that $\|T\| \leq W$. Thus

$$|f_0(a) - f_0(b)| \leq 2W\|a-b\|.$$

Lemma 4.21. *Let $\mathscr{H}$ be a Hilbert space over $\mathbf{D}$ of finite dimension n. If f is a nonnegative frame function which is regular on each completely real subspace, then f is regular.*

Proof. We prove this by induction on n. For $n=1$, the lemma is trivial. Let $n>1$. Let $d = \sup_{\|u\|=1} f(u)$ and $\{x_n\}$ a sequence of points on $\mathscr{E}$, the unit sphere of $\mathscr{H}$, such that $f(x_n) \to d$. We may assume, in view of the compactness of the unit sphere, that $x_n \to y$, $\|y\|=1$. We first show that $f(y)=d$. Let $\lambda_n = |\langle y, x_n \rangle|^{-1} \langle y, x_n \rangle$. Since $x_n \to y$, $\lambda_n \to 1$. The inner product $\langle \lambda_n x_n, y \rangle$ is real and consequently $\lambda_n x_n$ and y span, over $\mathbf{R}$, a completely real subspace. We now have, if W is the weight of f,

$$\begin{aligned} |f(y)-d| &\leq |f(y) - f(\lambda_n x_n)| + |f(\lambda_n x_n) - d| \\ &\leq 2W\|y - \lambda_n x_n\| + |f(x_n) - d|, \end{aligned}$$

since f, by hypothesis, is regular when restricted to the completely real subspace spanned by $\lambda_n x_n$, y, and $f(x_n) = f(\lambda_n x_n)$ $(|\lambda_n|=1)$. This shows that $f(y)=d$.

Let us now extend the function f to a function F defined on all of $\mathscr{H}$ by setting

$$F(v) = \begin{cases} 0 & v = 0. \\ \|v\|^2 f(\|v\|^{-1} v), & v \neq 0 \end{cases} \tag{52}$$

Clearly for $c \in \mathbf{D}$,

$$F(cv) = |c|^2 F(v). \tag{53}$$

We now assert that if u is any vector of $\mathscr{H}$ orthogonal to y, then

$$F(cy+u) = |c|^2 d + F(u) \tag{54}$$

for all $c \in \mathbf{D}$. Let L be the two-dimensional real linear manifold spanned by y and u. This is completely real. f is regular on its unit sphere so that f is the restriction to L of a nonnegative definite quadratic form Q. But y is the point at which the maximum of the quadratic form is reached and u is orthogonal to y in L. Thus the matrix of the quadratic form Q is diagonal with respect to the basis $\{y, u\}$ of L, and consequently we have, whenever r and r' are two *real* numbers,

$$F(ry + r'u) = |r|^2 d + |r'|^2 F(u).$$

It is to be noted that u here is an arbitrary vector orthogonal to y.

Suppose now that v is any vector orthogonal to y and $c \neq 0$ an element from $\mathbf{D}$, not necessarily real. We have:

$$\begin{aligned} F(cy+v) &= F(c(y+c^{-1}v)) \\ &= |c|^2 F(y+c^{-1}v) \\ &= |c|^2 F(y+|c|^{-1}|c|c^{-1}v), \end{aligned}$$

but $u=|c|c^{-1}v$ is a vector orthogonal to y. Thus

$$\begin{aligned} F(y+c^{-1}v) &= d+|c|^{-2}F(|c|c^{-1}v) \\ &= d+|c|^{-2}F(v), \end{aligned}$$

proving that

$$F(cy+v) = |c|^2 d + F(v).$$

The restriction of f to the orthogonal complement $\mathscr{K}$ of y is a frame function, which, on $\mathscr{K}$, has the same properties as f. Suppose we assume the present lemma to be true for all Hilbert spaces of dimension $k<n$. Then F restricted to $\mathscr{K}$ is regular by the induction hypothesis and hence there is an orthonormal basis $\{e_1, \cdots, e_{n-1}\}$ for $\mathscr{K}$ and real numbers $d_1, \cdots, d_{n-1} \geq 0$ such that

$$F(c_1e_1+\cdots+c_{n-1}e_{n-1}) = |c_1|^2 d_1+\cdots+|c_{n-1}|^2 d_{n-1}$$

for all $c_1, \cdots, c_{n-1} \in \mathbf{D}$. By the proof given in the previous paragraph, this means that

$$F(cy+c_1e_1+\cdots+c_{n-1}e_{n-1}) = |c|^2 d+|c_1|^2 d_1+\cdots+|c_{n-1}|^2 d_{n-1},$$

which shows that f is the restriction to the unit sphere of the regular frame function f_U, where U is the operator defined by $Uy=dy$, $Ue_j=d_je_j$, $j=1, 2, \cdots, n-1$. This proves the lemma.

Lemma 4.22. *Let $\mathscr{H}$ be a separable Hilbert space over* $\mathbf{D}$ *which is one of* $\mathbf{R}$, $\mathbf{C}$, *or* $\mathbf{H}$ *of dimension at least* 3 *and f a nonnegative frame function on the unit sphere $\mathscr{E}$ of $\mathscr{H}$. Then f is regular.*

Proof. Since the dimension of $\mathscr{H}$ is at least 3, it follows that the restriction of f to any completely real subspace is regular. Only the case $\dim \mathscr{H}=\infty$ remains. Lemma 4.21 is applicable to any finite dimensional subspace of $\mathscr{H}$. Thus for any finite dimensional subspace S of $\mathscr{H}$ there is an operator U_S which is an endomorphism of S, bounded, self-adjoint, and ≥ 0 such that

$$f(x) = \langle U_S x,x\rangle \qquad (x \in S, \quad \|x\| = 1).$$

The uniqueness of U_S implies that if $S_1 \subseteq S_2$, $\langle U_{S_1}x,y\rangle=\langle U_{S_2}x,y\rangle$ if $x,y \in S_1$. This implies in the usual way that there exists a symmetric semi-bilinear form on $\mathscr{H} \times \mathscr{H}$, say, B such that

$$B(x,x) = f(x)$$

for all x with $\|x\|=1$. Since $0 \leq B(x,x) \leq 1$, there exists a bounded, non-negative, self-adjoint operator U such that

$$\langle Ux,x\rangle = B(x,x) = f(x).$$

This proves the lemma.

Theorem 4.23 (*Gleason* [1]). *Let $\mathscr{W}$ be the convex set of all von Neumann operators of unit trace on a separable Hilbert space $\mathscr{H}$ of dimension at least* 3 (*possibly infinite*) *over* $\mathbf{D}$. *Then the map*

$$U \to p_U \qquad (U \in \mathscr{W}) \tag{55}$$

(*where p_U is defined by* (30)) *is a convex*† *isomorphism of $\mathscr{W}$ onto the convex set $\mathscr{S}$ of all states of the logic $\mathscr{L}$. A state $p \in \mathscr{S}$ is pure if and only if $p=p_u$ for a unit vector u of $\mathscr{H}$, p_u being defined as in* (25). *$p_u=p_v$ if and only if for some $c \in \mathbf{D}$ with $|c|=1$, $v=cu$. In particular, the pure states of $\mathscr{L}$ are in natural one-one correspondence with the rays of $\mathscr{H}$.*

Proof. That $U \to p_U$ is an isomorphism of $\mathscr{W}$ onto $\mathscr{S}$ is immediate from lemma 4.22. Next we shall obtain a description of the pure states of $\mathscr{L}$. For any vector $u \in \mathscr{H}$ with $\|u\|=1$, let p_u be the state of $\mathscr{L}$ defined by

$$p_u(M) = \|P^M u\|^2 \qquad (M \in \mathscr{L}).$$

We assert that p_u is a pure state. Let U_0 be the projection on the one-dimensional space $\mathbf{D} \cdot u$. Clearly $p_u(M) = \operatorname{tr}(P^M U_0)$ so that

$$p_u = p_{U_0}.$$

To prove that p_u is pure, let p_1, p_2 be states of $\mathscr{L}$ such that

$$p_u = ap_1 + (1-a)p_2, \tag{56}$$

where $0 < a < 1$. Since $U \to P_U$ is a one-one convex map of $\mathscr{W}$ *onto* the set of all states, there are (unique) elements U_1, $U_2 \in \mathscr{W}$ such that $p_j = p_{U_j}$ $(j=1,2)$ and hence

$$U_0 = aU_1 + (1-a)U_2. \tag{57}$$

Since $\langle U_j x,x\rangle \geq 0$ for all $x \in \mathscr{H}$, $j=0$, 1, 2, and since $\langle U_0 x,x\rangle = 0$ if x is orthogonal to u, we conclude from (57) that for $x \perp u$, $\langle U_j x,x\rangle = 0$ $(j=1, 2)$. Now U_j is ≥ 0 and hence there is a bounded self-adjoint V_j such that $V_j^2 = U_j$ and we have $\langle U_j x,x\rangle = \|V_j x\|^2$. This implies that $V_j x = 0$ for all x orthogonal to u and, consequently, that $U_j x = 0$ for all x orthogonal to u. Since U_j is self-adjoint, this implies that $U_j = c_j U_0$ for some constant

---

† If A and B are convex subsets of two vector spaces over the real number field and if u is a map of A into B, we say that u is convex if $u(ax+by) = au(x) + bu(y)$ for $x, y \in A$, $0 \leq a, b \leq 1$ and $a+b=1$.

$c_j (j=1,2)$, and $c_j=1$ as $\operatorname{tr}(U_j)=1$. Therefore $U_1=U_2=U_0$, which proves that p_u is a pure state. Conversely let $U_0 \in \mathscr{W}$ be such that p_{U_0} is a pure state. If $a_1, \cdots, a_n, \cdots$ $(\geq 0, a_1+a_2+\cdots=1)$ are the eigenvalues of U_0, and $u_1, u_2, \cdots, u_n, \cdots$ corresponding eigenvectors of norm 1, then

$$p_{U_0} = \sum_n a_n p_{u_n}. \tag{58}$$

As p_{U_0} is pure, (58) implies that all but one of the a_n must be 1, the rest being 0. This shows that $p_{U_0}=p_{u_j}$ for some j.

Remark. From theorem 4.23 we can obtain at once an important characterization of the principle of superposition of pure states. Let $\{u_\alpha\}$ be a family of unit vectors of $\mathscr{H}$ and let us write p_α for p_{u_α}. Let $p=p_u$ be a pure state and let M be the smallest closed linear manifold containing all the u_α. If p is a superposition of the collection $\{p_\alpha\}$, then $p(M^\perp)=0$ as $p_\alpha(M^\perp)=0$ for all α. Hence $u \in M$. On the other hand, if $u \in M$, then for any $N \in \mathscr{L}$ having the property that $p_\alpha(N)=0$ for all α, $p(N)=0$, since any such N is contained in $M^\perp$. In other words, p_u is a superposition of the family $\{p_{u_\alpha}\}$ if and only if u lies in the smallest closed linear manifold of $\mathscr{H}$ containing all the u_α. Thus, if we make the assumption that the logic of the physical system $\mathfrak{S}$ is standard, *the pure states can be identified with the points of an infinite dimensional projective geometry with the principle of superposition translating into the usual concept of linear dependence of points.*

Our next object is to discuss the nature of the statistics of one or more observables associated with the standard logic $\mathscr{L}$. We begin with the following theorem:

Theorem 4.24. *Let $x(E \rightarrow x(E))$ be an observable associated with $\mathscr{L}$, U a von Neumann operator on $\mathscr{H}$ of trace 1 and p_U, the state $M \rightarrow \operatorname{tr}(P^M U)$ of $\mathscr{L}$. Let A_x be the self-adjoint operator on $\mathscr{H}$ which corresponds to x. If f is any real Borel function bounded on the spectrum of x, $f \circ x$ is a bounded observable and*

$$\mathscr{E}(f \circ x | p_U) = \operatorname{tr}(f(A_x)U). \tag{59}$$

If x is nonnegative, so is A_x and then $\mathscr{E}(x|p_U)$ exists if and only if $B=A_x^{1/2}U^{1/2}$ is an everywhere defined, bounded operator of Hilbert-Schmidt class; in this case

$$\mathscr{E}(x|p_U) = \operatorname{tr}(B^*B). \tag{60}$$

If x is arbitrary, then x has finite variance in the state p_U if and only if $A_x U^{1/2}$ is an everywhere defined, bounded operator of Hilbert-Schmidt class. In this case, (60) *reduces to*

$$\mathscr{E}(x|p_U) = \operatorname{tr}(U^{1/2}A_x U^{1/2}). \tag{61}$$

In particular, if u *is a unit vector in* $\mathscr{H}$ *and* p_u *is the state* $M \to \|P^M u\|^2$ *of* $\mathscr{L}$, x *has a variance in the state* p_u *if and only if* $u \in \mathscr{D}(A_x)$;† *in this case,*

$$\mathscr{E}(x|p_u) = \langle A_x u, u \rangle. \tag{62}$$

Proof. If S is the spectrum of x and $K = (f[S])^-$, the closure of $f[S]$, then $(f \circ x)(K) = \mathscr{H}$; and as K is compact, this proves that $y = f \circ x$ is bounded. The spectral measure of the operator A_y is the measure $E \to Q_{f^{-1}(E)}$ where Q is the spectral measure of A_x. Thus $A_y = f(A_x)$ in the usual sense of the functional calculus of self-adjoint operator on $\mathscr{H}$. This proves (59) at once.

We next take up the more delicate questions involving possibly unbounded x. Let x be any nonnegative observable i.e., $x([0,\infty[) = \mathscr{H}$, associated with $\mathscr{L}$. Then A_x is a nonnegative, self-adjoint operator. We denote by $A_x^{1/2}$ the unique, nonnegative, self-adjoint operator C such that $C^2 = A_x$. Let U be any von Neumann operator of unit trace on $\mathscr{H}$ such that $\mathscr{E}(x|p_U) < \infty$. Let $Q(E \to Q_E)$ be the spectral measure of A_x and $k(E) = \mathrm{tr}(Q_E U)$. Clearly $k([0,\infty[) = 1$ and

$$\mathscr{E}(x|p_U) = \int_0^\infty t\,dk(t).$$

If $\varphi \in \mathscr{H}$ is any unit vector, $\langle Q_E U^{1/2}\varphi, U^{1/2}\varphi \rangle \leq \mathrm{tr}(U^{1/2} Q_E U^{1/2}) = k(E)$; hence, on writing $\psi = U^{1/2}\varphi$,

$$\int_0^\infty (t^{1/2})^2 dk_\psi(t) < \infty,$$

where k_ψ is the measure $E \to \langle Q_E \psi, \psi \rangle$. This implies that ψ lies in the domain of $A_x^{1/2}$. In other words, $A_x^{1/2} U^{1/2}$ is everywhere defined. Since $A_x^{1/2}$ is closed and $U^{1/2}$ is bounded, it is easily seen that $A_x^{1/2} U^{1/2}$ is closed. Consequently, $B = A_x^{1/2} U^{1/2}$ is a bounded operator. To prove that B is of Hilbert-Schmidt class it is enough to prove that $\sum_n \|B\varphi_n\|^2 < \infty$ for some orthonormal basis of $\mathscr{H}$. Let $\{\varphi_n\}$ be an orthonormal basis of $\mathscr{H}$ such that $U\varphi_n = a_n \varphi_n$ for all n; then $a_n \geq 0$, $\sum_n a_n = 1$. Let $k_n(E) = \langle Q_E \varphi_n, \varphi_n \rangle$. Then each k_n is a probability measure and $k = \sum_n a_n k_n$. But

$$\begin{aligned}
\sum_n \|B\varphi_n\|^2 &= \sum_n a_n \|A_x^{1/2}\varphi_n\|^2 \\
&= \sum_n a_n \int_0^\infty (t^{1/2})^2\, dk_n(t) \\
&= \sum_n a_n \int_0^\infty t\,dk_n(t) \\
&= \int_0^\infty t\,dk(t) \\
&< \infty.
\end{aligned}$$

---

† $\mathscr{D}(L)$ is the domain of the linear transformation L.

For the converse, if $B = A_x^{1/2}U^{1/2}$ is defined everywhere and is a bounded Hilbert-Schmidt operator, we have, using the above notation,

$$\mathrm{tr}(B^*B) = \sum_n \|B\varphi_n\|^2 = \sum_n a_n \int_0^\infty t\,dk_n(t) < \infty.$$

But then, by Fatou's lemma,

$$\int_0^\infty t\,dk(t) = \mathscr{E}(x|p_U) < \infty.$$

At the same time, the above argument proves that

$$\mathscr{E}(x|p_U) = \mathrm{tr}(B^*B).$$

Next we consider the case of finite variance. We use the same notation. x has a finite variance in the state p_U if and only if $\int_{-\infty}^{\infty} t^2 dk(t) < \infty$. Proceeding exactly as before, we conclude that if this condition is satisfied, $D = A_x U^{1/2}$ is an everywhere defined operator of Hilbert-Schmidt class. Note that, if $a_n > 0$, $\varphi_n = a_n^{-1/2}U^{1/2}\varphi_n$ so that $\varphi_n \in \mathscr{D}(A_x)$. In this case, $\langle A_x\varphi_n, \varphi_n\rangle = \int_{-\infty}^{\infty} t\,dk_n(t)$ by standard spectral theory. Now, $U^{1/2}D$ is a bounded operator and, being symmetric, is self-adjoint. As $U^{1/2}$ is also Hilbert-Schmidt, $U^{1/2}D$ is of trace class. We then obtain

$$\begin{aligned}
\mathrm{tr}(U^{1/2}D) &= \sum_n \langle A_x U^{1/2}\varphi_n, U^{1/2}\varphi_n\rangle \\
&= \sum_{n:a_n>0} a_n\langle A_x\varphi_n, \varphi_n\rangle \\
&= \sum_n a_n \int_{-\infty}^{\infty} t\,dk_n(t) \\
&= \int_{-\infty}^{\infty} t\,dk(t).
\end{aligned}$$

This proves (61). For proving the converse, let $A_x U^{1/2} = D$ be everywhere defined and of Hilbert-Schmidt class. We want to prove that

$$\int_{-\infty}^{\infty} t^2 dk(t) < \infty.$$

Write, for any Borel set $E \subseteq R^1$,

$$k_n(E) = \langle Q_E U^{1/2}\varphi_n, U^{1/2}\varphi_n\rangle.$$

Then

$$\begin{aligned}
k(E) &= \mathrm{tr}(U^{1/2}Q_E U^{1/2}) \\
&= \sum_n k_n(E)
\end{aligned}$$

and

$$\begin{aligned}
\sum_n \int_{-\infty}^{\infty} t^2 dk_n(t) &= \sum_n \|A_x U^{1/2}\varphi_n\|^2 \\
&< \infty.
\end{aligned}$$

It follows once again by Fatou's lemma that $\int_{-\infty}^{\infty} t^2 dk(t) < \infty$.

For the last part, let U be the projection on the one-dimensional space spanned by φ. Then $U^{1/2}=U$ and $A_xU^{1/2}$ is defined everywhere if and only if $\varphi \in \mathscr{D}(A_x)$; in this case, it is easily seen that $A_xU^{1/2}$ has at most one-dimensional range and $\operatorname{tr}(U^{1/2}A_xU^{1/2})=\langle A_x\varphi,\varphi\rangle$ as $\|\varphi\|=1$. Theorem 4.24 is completely proved.

We next come to the analysis of the joint statistics of several observables. For simplicity we deal only with *bounded* observables. Let x_j $(1\leq j\leq k)$ be bounded observables associated with $\mathscr{L}$. Let A_j be the bounded self-adjoint operator which corresponds to x_j. If $c_1, c_2, \cdots, c_k \in \mathbf{D}$ are *real* numbers, and $\mathbf{c}=(c_1, \cdots, c_k)$, $A(\mathbf{c})=c_1A_1+\cdots+c_kA_k$ is also a bounded self-adjoint operator. Let $E \rightarrow Q(\mathbf{c}; E)$ be the spectral measure of $A(\mathbf{c})$. If $\varphi \in \mathscr{H}$ is a unit vector, and if we define $q_\varphi(\mathbf{c}; E)$ by

$$q_\varphi(\mathbf{c}; E) = \|Q(\mathbf{c}; E)\varphi\|^2, \tag{63}$$

$q_\varphi(\mathbf{c};\cdot)$ is a probability measure on $\mathscr{B}(\mathbf{R}^1)$. We say that the observables $\{x_1, \cdots, x_k\}$ have *joint statistical distributions* if for each unit vector $\varphi\in\mathscr{H}$, there exists a probability measure q_φ on $\mathscr{B}(\mathbf{R}^k)$ such that for all vectors $\mathbf{c}=(c_1,\cdots,c_k)$ of $\mathbf{R}^k$ and all $E\in\mathscr{B}(\mathbf{R}^1)$,

$$q_\varphi(E(\mathbf{c})) = q_\varphi(\mathbf{c}; E), \tag{64}$$

where

$$E(\mathbf{c}) = \{(t_1, \cdots, t_k) : c_1t_1+\cdots+c_kt_k \in E\}. \tag{65}$$

The sets $E(\mathbf{c})$ are the "half-spaces" in $\mathbf{R}^k$. Since any $\mathbf{D}$-valued measure on $\mathscr{B}(\mathbf{R}^k)$ is uniquely determined by its values on the sets of the form $E(\mathbf{c})$, it follows that q_φ, if it exists at all, is unique. Obviously, the above sense is a quite conservative one in which we may speak of the statistics of the observables $\{x_1, \cdots, x_k\}$; the probability measure q_φ will then be the joint distribution of $\{x_1,\cdots, x_k\}$ under the state of $\mathscr{L}$ determined by φ. We shall now prove the following theorem (Varadarajan [1]).

Theorem 4.25. *With the notations described above, $\{x_1, \cdots, x_k\}$ have joint statistical distributions if and only if the operators $A_j(1\leq j\leq k)$ commute with one another.*

Proof. If the A_j commute with one another, the observables x_j are simultaneously observable and hence there exists, by theorem 3.17, a σ-homomorphism β of $\mathscr{B}(\mathbf{R}^k)$ into $\mathscr{L}$ such that

$$\beta(\pi_j{}^{-1}(E)) = x_j(E) \tag{66}$$

for all $E\in\mathscr{B}(\mathbf{R}^1)$ and $j=1, 2, \cdots, k$, π_j being the projection $(t_1, \cdots, t_k) \rightarrow t_j$ of $\mathbf{R}^k$ into $\mathbf{R}^1$. If we define for a unit vector $\varphi\in\mathscr{H}$, and $E\in\mathscr{B}(\mathbf{R}^k)$,

$$q_\varphi(E) = \|P^{\beta(E)}\varphi\|^2, \tag{67}$$

then (63) and (64) are satisfied. Thus $\{x_1, \cdots, x_k\}$ possess joint distributions.

We now proceed to prove the converse. If $\{x_1, \cdots, x_k\}$ possess joint distributions, it is obvious that for $i, j \leq k$, $i \neq j$, $\{x_i, x_j\}$ have joint distributions. We may (and do) therefore assume that $k=2$ and prove that $A_1A_2 = A_2A_1$. For any $\varphi \in \mathscr{H}$ which is nonzero, we write $\psi = (\|\varphi\|)^{-1}\varphi$ and define $q_\varphi = \|\varphi\|^2 q_\psi \cdot q_\varphi$ is a nonnegative measure on $\mathscr{B}(\mathbf{R}^2)$ and $q_\varphi(\mathbf{R}^2) = \|\varphi\|^2$. Set $q_0 = 0$. We now define the $\mathbf{D}$-valued measure $q_{\varphi,\varphi'}$, for an arbitrary pair of vectors φ, φ', as follows:

$$(68)\qquad \begin{aligned} q_{\varphi,\varphi'} &= \tfrac{1}{4}(q_{\varphi+\varphi'} - q_{\varphi-\varphi'}) \qquad (\mathbf{D} = \mathbf{R}), \\ q_{\varphi,\varphi'} &= \tfrac{1}{4}(q_{\varphi+\varphi'} - q_{\varphi-\varphi'} + iq_{-i\varphi+\varphi'} - iq_{-i\varphi-\varphi'}) \qquad (\mathbf{D} = \mathbf{C}), \\ q_{\varphi,\varphi'} &= \tfrac{1}{4}\left(q_{\varphi+\varphi'} - q_{\varphi-\varphi'} + \sum_{r=1}^{3} (q_{-\mathbf{j}_r\varphi+\varphi'} - q_{-\mathbf{j}_r\varphi-\varphi'})\mathbf{j}_r\right) \qquad (\mathbf{D} = \mathbf{Q}). \end{aligned}$$

We assert that for each $M \in \mathscr{B}(\mathbf{R}^2)$, the map $\varphi, \varphi' \to q_{\varphi,\varphi'}(M)$ is bounded and $*$-bilinear on $\mathscr{H} \times \mathscr{H}$. To verify this, let φ_1, φ_2, $\varphi' \in \mathscr{H}$ and let us write

$$\nu(M) = q_{\varphi_1+\varphi_2,\varphi'}(M) - q_{\varphi_1,\varphi'}(M) - q_{\varphi_2,\varphi'}(M);$$

then ν is a $\mathbf{D}$-valued measure on $\mathscr{B}(\mathbf{R}^2)$. However, if $M = E(\mathbf{c})$, where $E \in \mathscr{B}(\mathbf{R}^1)$, then $q_\varphi(M) = \langle Q(\mathbf{c};\, E)\varphi, \varphi\rangle$ and hence from (68) we infer that

$$q_{\varphi,\varphi'}(E(\mathbf{c})) = \langle Q(\mathbf{c};\, E)\varphi, \varphi'\rangle.$$

Consequently, $\nu(M) = 0$ whenever $M = E(\mathbf{c})$. Hence $\nu = 0$. Similar arguments complete the proof that $\varphi, \varphi' \to q_{\varphi,\varphi'}(M)$ is $*$-bilinear. Since

$$0 \leq q_{\varphi,\varphi}(M) \leq \|\varphi\|^2,$$

it is clear that $\varphi, \varphi' \to q_{\varphi,\varphi'}(M)$ is a bounded form. We can now use a well known theorem to construct a unique bounded self-adjoint operator P_M such that

$$q_{\varphi,\varphi'}(M) = \langle P_M\varphi, \varphi'\rangle.$$

Clearly, $0 \leq P_M \leq I$ for all M, $P_\phi = 0$, and $P_{R^2} = I$. Moreover, as $M \to q_{\varphi,\varphi'}(M)$ is a measure, P_M is additive over disjoint sets M, i.e., if $M_1, M_2, \cdots$ are pairwise disjoint Borel sets of $\mathbf{R}^2$ and $M = \bigcup_n M_n$,

$$\langle P_M\varphi, \varphi'\rangle = \sum_{n=1}^{\infty} \langle P_{M_n}\varphi, \varphi'\rangle.$$

We note also, as $q_{\varphi,\varphi}(M) \geq 0$, that for $M_1 \subseteq M_2$, $P_{M_1} \leq P_{M_2}$. Finally, if $M = E(\mathbf{c})$, $P_M = Q(\mathbf{c};\, E)$. To complete the proof we need another lemma.

Lemma 4.26. *Let B be a nonnegative self-adjoint operator and S, T two projections such that $B \leq S$ and $B \leq T$. Then $B \leq S \wedge T$.*

Proof. Let X and Y be the ranges of S and T, respectively, and let $C = B^{1/2}$. Since $\|Cx\|^2 \leq \|Sx\|^2$, C, and hence B, vanishes on $X^\perp$. Similarly,

B vanishes on $Y^\perp$ and hence $B=0$ on $(X \cap Y)^\perp \cdot B$, being self-adjoint, therefore leaves $X \cap Y$ invariant and for $x \in X \cap Y$, $\langle Bx,x\rangle \leq \|x\|^2$. Thus $B \leq P^{X \cap Y} = S \wedge T$.

We are now in a position to complete the argument. We must show that for any two sets E_1, E_2 in $\mathscr{B}(\mathbf{R}^1)$, $Q((1,0);E_1)$ and $Q((0,1);E_2)$ commute. Write $Q_1=Q((1,0);E_1)$ and $Q_2=Q((0,1);E_2)$. Since $E_1 \times E_2 \subseteq E_1 \times \mathbf{R}^1$, we have

$$P_{E_1 \times E_2} \leq P_{E_1 \times R^1} = Q_1.$$

Similarly,

$$P_{E_1 \times E_2} \leq Q_2.$$

From lemma 4.26 we may conclude that

$$P_{E_1 \times E_2} \leq Q_1 \wedge Q_2.$$

Replacing E_2 by $\mathbf{R}^1 - E_2$ we obtain

$$P_{E_1 \times R^1 - E_2} \leq Q_1 \wedge (I-Q_2).$$

The last two relations yield, on addition,

$$Q_1 \leq Q_1 \wedge Q_2 + Q_1 \wedge (I-Q_2). \tag{69}$$

However, $Q_1 \wedge Q_2$ and $Q_1 \wedge (I-Q_2)$ are orthogonal projections which are both $\leq Q_1$. Hence we infer from (69) that

$$Q_1 = Q_1 \wedge Q_2 + Q_1 \wedge (I-Q_2).$$

This last equation leads quite simply to the conclusion that Q_1 and Q_2 commute (cf. lemma 3.7).

This theorem, taken together with theorem 3.9 may be regarded as giving a complete description of the circumstances under which the statistics of several observables may be regarded as arising from observables with joint distributions.

We end these remarks with a mention of the concept of the quasi-probability distributions, introduced by Wigner [2] and studied extensively by Moyal [1]. Let $x_1, x_2, \cdots, x_k$ be bounded observables with $A_1, \cdots, A_k$ as their corresponding operators. We use the same notation as in theorem 4.25. For fixed φ in $\mathscr{H}$, and for a given $\mathbf{c}$, $E \to Q(\mathbf{c};E) = q_\varphi(\mathbf{c};E)$ is a non-negative measure on $\mathscr{B}(\mathbf{R}^1)$. We define

(70)

$$\Phi(c_1, c_2, \cdots, c_k; \varphi) = \frac{1}{(2\pi)^{k/2}} \int_{\mathbf{R}^k} \exp\{(-1)^{1/2}(c_1t_1 + \cdots + c_kt_k)\} dq_\varphi(\mathbf{c};\cdot).$$

It is easy to prove that $\Phi(\cdot;\varphi)$ is a continuous function of $c_1, \cdots, c_k$ for each $\varphi \in \mathscr{H}$ and

$$|\Phi(c_1, \cdots, c_k; \varphi)| \leq 1.$$

We may therefore introduce its Fourier transform, in the sense of distributions of Schwartz, say $\hat{\Phi}(\cdot\,;\varphi)$. It is known as the *Wigner-Moyal quasi-probability distribution* of the observables $x_1, \cdots, x_k$ in the state determined by φ.

3. THE STANDARD LOGICS: SYMMETRIES

At the focal point of the axiomatic development of modern physics lies the concept of automorphism or symmetry. We have already pointed out that the momentum observables of a physical system are intimately related to one-parameter groups of automorphisms of the configuration manifold of the system. Moreover, we had described in Chapter III how the dynamical group of a physical system can be described by a one-parameter group of convex automorphisms of the set of states. It is therefore extremely important to examine the concept of an automorphism when the logic in question is a standard one. The main purpose of this section is to obtain a complete description of the automorphisms associated with a standard logic.

We shall introduce, for this purpose, the notion of a symmetry of a Hilbert space $\mathscr{H}$. As always, $\mathscr{H}$ is separable and the field of scalars $\mathbf{D}$ is one of $\mathbf{R}$, $\mathbf{C}$, or $\mathbf{H}$. A mapping $T(x \to Tx)$ of $\mathscr{H}$ into itself is said to be a *symmetry* (of $\mathscr{H}$) if (i) T is additive, one-one, and maps $\mathscr{H}$ onto itself, and (ii) there exists a continuous automorphism θ of $\mathbf{D}$ such that T is θ-linear and, for all $x, y \in \mathscr{H}$,

$$\langle Tx,Ty\rangle = \langle x,y\rangle^{\theta}. \tag{71}$$

If $\mathbf{D}=\mathbf{R}$, then θ is the identity and the symmetries are none other than the unitary operators of $\mathscr{H}$. If $\mathbf{D}=\mathbf{C}$, then θ is either the identity or the conjugation; in the former case, T is unitary, whereas in the latter case T is anti-unitary. If $\mathbf{D}=\mathbf{H}$, then there exists a $c \in \mathbf{D}$ such that $|c|=1$ and $d^{\theta}=cdc^{-1}$ for all $d \in \mathbf{D}$; in this case,

$$T(dx) = (cdc^{-1})Tx$$

and

$$\langle Tx,Ty\rangle = c\langle x,y\rangle c^{-1}.$$

Equation (71) implies that

$$\left.\begin{aligned}\|Tx\| &= \|x\| \\ |\langle Tx,Ty\rangle| &= |\langle x,y\rangle|\end{aligned}\right. \qquad (x, y \in \mathscr{H}).$$

In the general case of (71) we shall call T *θ-unitary*. The symmetries of $\mathscr{H}$ form a group; if T_j is θ_j-unitary $(j=1,2)$, T_1T_2 is $\theta_1\theta_2$-unitary and T_j^{-1} is θ_j^{-1}-unitary. The unitary (linear) operators of $\mathscr{H}$ form a normal subgroup

$\mathscr{U}(\mathscr{H})$ of the group of all symmetries. We shall write $S(\mathscr{H})$ for the group of symmetries. The map which associates with any θ-unitary T, the automorphism θ of $\mathbf{D}$, is a homomorphism of $S(\mathscr{H})$ onto the group $\mathrm{Aut}(\mathbf{D})$ of all continuous automorphisms of $\mathbf{D}$. Since the kernel of this map is $\mathscr{U}(\mathscr{H})$, we have:

$$S(\mathscr{H})/\mathscr{U}(\mathscr{H}) \simeq \mathrm{Aut}(\mathbf{D}). \tag{72}$$

Note finally that if T is a symmetry and $c \in \mathbf{D}$ is such that $|c| = 1$, cT is a symmetry.

Theorem 4.27. *Let $\mathscr{H}$ be a separable infinite dimensional Hilbert space over $\mathbf{D}$. Let $\alpha(M \to M^\alpha)$ be an automorphism of the logic $\mathscr{L}$. Then, there exists a symmetry T of $\mathscr{H}$ such that*

$$M^\alpha = T[M] \qquad (M \in \mathscr{L}). \tag{73}$$

If T' is a symmetry such that (73) *is satisfied for all M, then there exists a $c \in D$ with $|c| = 1$ such that $T' = cT$. Finally if T is any symmetry of $\mathscr{H}$, $M \to T[M]$ is an automorphism of the logic $\mathscr{L}$.*

Proof. It is quite trivial to check that if T is a symmetry of $\mathscr{H}$, $M \to T[M]$ is an automorphism of the logic $\mathscr{L}$.

Suppose now that $\alpha(M \to M^\alpha)$ is an automorphism of the logic $\mathscr{L}$. Let F be a linear manifold, chosen once for all, such that $\dim F = 3$. By theorem 2.1 there exists an automorphism θ of $\mathbf{D}$ and a θ-linear isomorphism L_0, of F onto F^α, such that for any linear manifold $M \subseteq F$,

$$M^\alpha = L_0[M].$$

Let us define $\mathscr{F}$ to be the collection of all finite dimensional linear manifolds G which contain F. We claim that for each $G \in \mathscr{F}$ there exists a unique θ-linear isomorphism L_G of G onto G^α such that

$$L_G u = L_0 u \qquad (u \in F) \tag{74}$$

and

$$M^\alpha = L_G[M] \qquad (M \subseteq G). \tag{75}$$

In fact, by theorem 2.1 there exists an automorphism θ' of $\mathbf{D}$ and a θ'-linear isomorphism L' of G onto G^α such that $M^\alpha = L'[M]$ for all linear manifolds $M \subseteq G$. Since $M^\alpha = L'[M] = L_0[M]$ for all linear manifolds $M \subseteq F$, we conclude from the same theorem that there exists a nonzero $c \in \mathbf{D}$ such that (i) $L'u = cL_0u$ for all $u \in F$, and (ii) $a^{\theta'} = ca^\theta c^{-1}$ for all $a \in \mathbf{D}$. If we write $L_G = c^{-1}L'$, then L_G is a θ-linear isomorphism of G onto G^α satisfying (74) and (75). The uniqueness of L_G follows from (74) and (75).

Since L_G is uniquely determined by (74) and (75), it follows quite easily that there exists a θ-linear map L of $\mathscr{H}$ into itself such that

$$Lu = L_G u \qquad (u \in G, \quad G \in \mathscr{F}). \tag{76}$$

We claim that L is one-one and maps $\mathscr{H}$ onto itself.

To prove that L is one-one, let $x \in \mathscr{H}$ and $Lx = 0$. Evidently, $x \in G$ for some $G \in \mathscr{F}$ and $Lx = L_G x = 0$. Therefore $x = 0$ as L_G is an isomorphism. If $y \in \mathscr{H}$ there is a $G \in \mathscr{F}$ such that $y \in G^\alpha$ and once again, as L_G maps G onto G^α, there will exist an $x \in G$ such that $Lx = y$. Since any $M \in \mathscr{L}$ with finite dimension is contained in some element of $\mathscr{F}$, we see that L is a θ-linear isomorphism of $\mathscr{H}$ onto itself such that

$$M^\alpha = L[M] \qquad (\dim M < \infty). \tag{77}$$

Let now $G \in \mathscr{F}$. Since α preserves orthogonal complementation, it follows that for $x, y \in G$, $\langle x,y \rangle = 0$ if and only if $\langle Lx,Ly \rangle = 0$. Thus, the semi-bilinear forms $x, y \to \langle x, y \rangle$ and $x, y \to \langle Lx,Ly \rangle^{(\theta-1)}$ of $G \times G$ induce the same polarity of the geometry of subspaces of G, and hence there exists, by virtue of theorem 2.3, a unique $d_G \neq 0$ in $\mathbf{D}$ such that

$$\langle Lx,Ly \rangle = \langle x,y \rangle^\theta d_G \qquad (x, y \in G).$$

Since $G \in \mathscr{F}$ is arbitrary and since d_G is uniquely determined for each G, it follows that d_G is independent of G and that

$$\langle Lx,Ly \rangle = \langle x,y \rangle^\theta d \qquad (x, y \in \mathscr{H}) \tag{78}$$

for a nonzero $d \in \mathbf{D}$.

We now show that θ is continuous. This is obvious for $\mathbf{D} = \mathbf{R}$ or $\mathbf{H}$, since θ is the identity in the former case and is an inner automorphism in the latter. Consider now the case $\mathbf{D} = \mathbf{C}$. If θ is not continuous, there exists a sequence $\{c_n\}$ in $\mathbf{D}$ such that $c_n \to 0$, but $|c_n{}^\theta| \to \infty$. Choose an orthonormal *infinite* sequence $\{e_n : n = 1, 2, \cdots\}$ in $\mathscr{H}$ and let $x_n = ne_n$. By replacing c_n by a subsequence if necessary, we may assume that

$$\|Lx_n\|/|c_n{}^\theta| \to 0 \qquad (n \to \infty). \tag{79}$$

Write

$$z = \sum_n \frac{c_n{}^*}{n} e_n.$$

Since $c_n{}^* \to 0$, z is well defined in $\mathscr{H}$. Clearly, $z \neq 0$. Then, $\langle x_n,z \rangle = c_n$ and hence, for all n, $u_n = \langle z,z \rangle^{-1} z - c_n{}^{-1} x_n$ is orthogonal to z. Consequently, from (78) we obtain, for all n,

$$\langle Lu_n,Lz \rangle = 0.$$

This implies, on simplification, that

$$(80) \qquad (\langle z,z\rangle^{\theta})^{-1}\langle Lz,Lz\rangle = (c_n{}^{\theta})^{-1}\langle Lx_n,Lz\rangle$$

for all n. But

$$|(c_n{}^{\theta})^{-1}||\langle Lx_n,Lz\rangle| \leq |c_n{}^{\theta}|^{-1}\|Lx_n\|\|Lz\|,$$

which tends to zero by virtue of our assumption (79). Hence (80) implies that $\langle Lz,Lz\rangle=0$ or $z=0$. This contradiction shows that θ is continuous.

We are now ready to complete the proof of theorem 4.27. Since θ is continuous, it follows that for any real number r, $r^{\theta}=r$. Hence we see, by putting $x=y$ in (78), that d is a positive real number. Write $T=d^{-1/2}L$. Then T is a θ-linear isomorphism of $\mathscr{H}$ onto itself such that

$$(81) \qquad \langle Tx,Ty\rangle = \langle x,y\rangle^{\theta}$$

and

$$M^{\alpha} = T[M] \qquad (M \in \mathscr{L}, \dim M < \infty).$$

Since $\|Tx\|=\|x\|$, T is an isometry and is hence continuous. It follows therefore that

$$M^{\alpha} = T[M] \qquad (M \in \mathscr{L}).$$

The remaining statements are easy consequences of theorem 2.3. Therefore, we omit their easy verifications. The proof of theorem 4.27 is complete.

We shall say that T *induces* α if (71) is satisfied.

Corollary 4.28. *If* $\mathbf{D}=\mathbf{R}$ *or* $\mathbf{H}$, *then any automorphism* α *of* $\mathscr{L}$ *is induced by a unitary operator of* $\mathscr{H}$. *If* $\mathbf{D}=\mathbf{C}$, *then* α^2 *is induced by a unitary operator.*

Proof. For $\mathbf{D}=\mathbf{R}$ there is nothing to prove. If $\mathbf{D}=\mathbf{H}$ and T_1 is a θ-unitary operator inducing α where θ is the automorphism

$$d \to cdc^{-1}(|c| = 1),$$

$T=c^{-1}T_1$ is a unitary operator inducing α. If $\mathbf{D}=\mathbf{C}$ and α is induced by a $*$-unitary T_1, $T_1{}^2$ is unitary and induces α^2.

In the physical literature it is customary to prove the above theorem in a somewhat different formulation (cf. Wigner [3]). To motivate Wigner's formulation we must introduce the notion of *transition probabilities*. Let $\mathbf{r}$, $\mathbf{r}'$ be rays in $\mathscr{H}$ and φ, φ' unit vectors in $\mathbf{r}$ and $\mathbf{r}'$, respectively. Then $|\langle\varphi,\varphi'\rangle|^2$ evidently depends only on $\mathbf{r}$ and $\mathbf{r}'$ and not on φ and φ'. We denote it by $[\mathbf{r},\mathbf{r}']$. It is called the *transition probability* between the *states* p and p' which correspond (cf. theorem 4.23) to the rays $\mathbf{r}$ and $\mathbf{r}'$. If a physical system $\mathfrak{S}$ is in the state p, $[\mathbf{r},\mathbf{r}']$ is the probability that it will be actually found in the state p' when an experiment is performed to decide whether it is in the state p' or not. We shall define a *physical symmetry associated with the logic* $\mathscr{L}$ to be any one-one map $\alpha(\mathbf{r} \to \mathbf{r}^\alpha)$ of the set of all rays of $\mathscr{H}$ onto itself with the property that

$$[\mathbf{r}^\alpha, \mathbf{r}'^\alpha] = [\mathbf{r}, \mathbf{r}'] \tag{82}$$

for all rays $\mathbf{r}$, $\mathbf{r}'$. The following theorem describes the natural connection between physical symmetries and symmetries of $\mathscr{H}$.

Theorem 4.29. (*Wigner* [3]). *Let $\mathscr{H}$ be a separable infinite dimensional Hilbert space over* $\mathbf{D}$ *and T a symmetry of $\mathscr{H}$. Then $\mathbf{r} \to T[\mathbf{r}]$ is a physical symmetry associated with $\mathscr{L}$. Conversely, let $\alpha(\mathbf{r} \to \mathbf{r}^\alpha)$ be any one-one map of the set $\mathscr{R}$ of all the rays of $\mathscr{H}$ onto itself with the property that $\mathbf{r} \perp \mathbf{r}'$ if and only if $\mathbf{r}^\alpha \perp \mathbf{r}'^\alpha$. Then, α is a physical symmetry associated with $\mathscr{L}$, and there exists a symmetry T of $\mathscr{H}$ such that $\mathbf{r}^\alpha = T[\mathbf{r}]$ for all $\mathbf{r} \in \mathscr{R}$. T is determined by α up to multiplication by a number $c \in \mathbf{D}$ with $|c| = 1$.*

Proof. If T is a symmetry of $\mathscr{H}$, $|\langle Tx, Ty\rangle| = |\langle x, y\rangle|$ for all $x, y \in \mathscr{H}$. From this we conclude easily that $\mathbf{r} \to T[\mathbf{r}]$ is a physical symmetry associated with $\mathscr{L}$. We now examine the converse. Let $\alpha(\mathbf{r} \to \mathbf{r}^\alpha)$ be any one-one map of $\mathscr{R}$ onto itself which preserves orthogonality. For any *subset* A of $\mathscr{R}$ let

$$A^\alpha = \{\mathbf{r}^\alpha : \mathbf{r} \in A\}, \tag{83}$$

$$[A] = \{x : x \in \mathscr{H}, x \in \mathbf{r} \quad \text{for some } \mathbf{r} \text{ in } A\}, \tag{84}$$

and

$$A^\perp = \{\mathbf{r} : \mathbf{r} \perp \mathbf{r}' \quad \text{for all} \quad \mathbf{r}' \in A\}. \tag{85}$$

We shall first prove that if $A \subseteq \mathscr{R}$ is such that $M = [A]$ is a closed linear manifold of $\mathscr{H}$, i.e., if A is the set of all rays of a closed linear manifold M of $\mathscr{H}$, then $[A^\alpha]$ is also a closed linear manifold. To prove this, we note that as $M = [A]$ is a closed linear manifold, a ray $\mathbf{r}$ belongs to A if and only if $\mathbf{r} \in (A^\perp)^\perp$. But as α preserves orthogonality, this means that a ray $\mathbf{r}' \in A^\alpha$ if and only if $\mathbf{r}' \in (A^{\perp\alpha})^\perp$, i.e., $A^\alpha = (A^{\perp\alpha})^\perp$. But for any *subset* $S \subseteq \mathscr{R}$, the set $[S^\perp]$ is a closed linear manifold. Consequently $[A^\alpha]$ is a closed linear manifold.

We now set up a correspondence, also denoted by α, between elements of $\mathscr{L}$. For any $M \in \mathscr{L}$, let A be the set of all rays $\mathbf{r}$ such that $\mathbf{r} \subseteq M$; we set:

$$M^\alpha = [A^\alpha]. \tag{86}$$

The argument of the preceding paragraph has shown that $M^\alpha \in \mathscr{L}$. By very definition it is clear that $\alpha(M \to M^\alpha)$ is an order-preserving map of $\mathscr{L}$. Since $\mathbf{r} \subseteq M^\perp$ if and only if $\mathbf{r}$ is orthogonal to all rays contained in M, it follows that $(M^\perp)^\alpha = (M^\alpha)^\perp$. Thus α preserves orthogonality. These results applied to the map α^{-1} of $\mathscr{R}$ show that the map $M \to M^{(\alpha^{-1})}$ defined by (86) is the inverse of the map $M \to M^{\alpha}$ of $\mathscr{L}$. Consequently, we may conclude that $M \to M^{\alpha}$ is a lattice automorphism of $\mathscr{L}$ which preserves orthogonality.

By theorem 4.27 there exists a symmetry T of $\mathscr{H}$ such that $M^{\alpha} = T[M]$ for all $M \in \mathscr{L}$. But then $\mathbf{r}^{\alpha} = T[\mathbf{r}]$ for all rays $\mathbf{r}$ showing that $\mathbf{r} \to \mathbf{r}^{\alpha}$ is a physical symmetry associated with $\mathscr{L}$. Since any automorphism of $\mathscr{L}$ is determined by its restriction to the set of all one-dimensional subspaces of $\mathscr{H}$, it is clear that T is determined by the map α of $\mathscr{R}$ up to multiplication by a number $c \in \mathbf{D}$ with $|c| = 1$. This completes the proof of the theorem.

In Chapter III we introduced the concept of a convex automorphism of the set of all states. We shall describe all such automorphisms of the set of states of a standard logic.

Let T be a symmetry of the Hilbert space $\mathscr{H}$, and $\mathscr{W}$ the convex set of all bounded, nonnegative, self-adjoint operators U of trace class with $\mathrm{tr}(U) = 1$. For any $U \in \mathscr{W}$ let $\alpha_T(U)$ be defined by

$$\alpha_T(U) = TUT^{-1}. \tag{87}$$

Lemma 4.30. *Let $U \in \mathscr{W}$ and let T be a symmetry of $\mathscr{H}$. Then $\alpha_T(U) \in \mathscr{W}$ and $U \to \alpha_T(U)$ is a convex automorphism of $\mathscr{W}$. Moreover, $T \to \alpha_T$ is a homomorphism of the group of symmetries of $\mathscr{H}$ into the group of convex automorphisms of $\mathscr{W}$.*

The proof is a routine verification. We leave it to the reader.

If T is a symmetry of $\mathscr{H}$ and if we set, for $U \in \mathscr{W}$,

$$\alpha_T(p_U) = p_{\alpha_T(U)}, \tag{88}$$

where p_U is the state $M \to \mathrm{tr}(P^M U)$ of $\mathscr{L}$, then the map $p_U \to \alpha_T(p_U)$ is a convex automorphism of $\mathscr{S}$. We shall prove that any convex automorphism of $\mathscr{S}$ is of this form. Let $\mathscr{W}_+$ denote the set of all von Neumann operators on $\mathscr{H}$. Clearly, $\mathscr{W} \subseteq \mathscr{W}_+$. If $T_1, T_2 \in \mathscr{W}_+$, we write $T_1 \leq T_2$ to mean $T_2 - T_1 \geq 0$.

Lemma 4.31. *Let α be a convex automorphism of $\mathscr{W}$. Then there exists a unique one-one map $\alpha^{\sim}$ of $\mathscr{W}_+$ onto itself such that*

$$(89)\quad \begin{aligned} &\text{(i)} \quad \alpha(U) = \alpha^{\sim}(U) \qquad (U \in \mathscr{W}),\\ &\text{(ii)} \quad \alpha^{\sim}(c_1A_1 + c_2A_2) = c_1\alpha^{\sim}(A_1) + c_2\alpha^{\sim}(A_2)\\ &\qquad\qquad (A_1, A_2 \in \mathscr{W}_+,\ c_1, c_2 \text{ real numbers} \geq 0); \end{aligned}$$

$\alpha^{\sim}$ has then the property:

$$(89)\quad \begin{aligned} &\text{(iii)} \quad \mathrm{tr}(\alpha^{\sim}(A)) = \mathrm{tr}(A) \qquad (A \in \mathscr{W}_+),\\ &\text{(iv)} \quad T_1 \leq T_2 \ \textit{ if and only if } \ \alpha^{\sim}(T_1) \leq \alpha^{\sim}(T_2). \end{aligned}$$

Proof. Define $\alpha^{\sim}(0)=0$. If $T \in \mathscr{W}_+$ and $T\neq 0$, $t=\mathrm{tr}(T)>0$, and $t^{-1}T \in \mathscr{W}$; we then put

$$\alpha^{\sim}(T) = t\alpha(t^{-1}T). \tag{90}$$

It is trivial to check that $\alpha^{\sim}$ has the properties described in the lemma. The uniqueness of $\alpha^{\sim}$ follows from the fact that the property (ii) of $\alpha^{\sim}$ implies (90). Equation (90) implies that $\mathrm{tr}(\alpha^{\sim}(T))=\mathrm{tr}(T)$, proving (iii). Finally, let $T_1 \leq T_2$. Then

$$\alpha^{\sim}(T_2) = \alpha^{\sim}(T_1)+\alpha^{\sim}(T_2-T_1) \geq \alpha^{\sim}(T_1)$$

as $\alpha^{\sim}(T_2-T_1) \in \mathscr{W}_+$ and is hence ≥ 0. Conversely, let $\alpha^{\sim}(T_1)\leq \alpha^{\sim}(T_2)$. Since $\alpha^{\sim}$ maps $\mathscr{W}_+$ onto itself, $\alpha^{\sim}(T_2)=\alpha^{\sim}(T_1)+\alpha^{\sim}(T)$ for some $T \in \mathscr{W}_+$. Evidently, $T_2=T_1+T$, and hence $T_2\geq T_1$.

Lemma 4.32. *Let $\varphi_j \in \mathscr{H}$ $(j=1, 2)$ be two unit vectors and let P_j be the orthogonal projection on the one-dimensional subspace $\mathbf{D}\cdot\varphi_j$. Then the two following statements are equivalent.*

(91)
(i) $\varphi_1 \perp \varphi_2$.
(ii) *For any $A \in \mathscr{W}_+$, the relations $P_1 \leq A$ and $P_2 \leq A$ imply that* $\mathrm{tr}\, A \geq 2$.

Proof. Suppose that $\varphi_1 \perp \varphi_2$ and suppose that $A \in \mathscr{W}_+$ satisfies the relations $P_1 \leq A$ and $P_2 \leq A$. Let $\varphi_3, \varphi_4, \cdots$ be vectors in $\mathscr{H}$ such that $\{\varphi_1, \varphi_2, \cdots\}$ is an orthonormal basis for $\mathscr{H}$. Then, as $A\geq 0$,

$$\begin{aligned}\mathrm{tr}(A) &= \sum_{n=1}^{\infty} \langle A\varphi_n,\varphi_n\rangle \\ &\geq \langle A\varphi_1,\varphi_1\rangle+\langle A\varphi_2,\varphi_2\rangle \\ &\geq \langle P_1\varphi_1,\varphi_1\rangle+\langle P_2\varphi_2,\varphi_2\rangle.\end{aligned}$$

Consequently,

$$\mathrm{tr}(A) \geq 2.$$

This shows that (i) implies (ii).

Conversely, let (ii) be satisfied. Let us assume that φ_1 is not orthogonal to φ_2. We shall construct an $A \in \mathscr{W}_+$ such that $P_1 \leq A$ and $P_2 \leq A$ but $\mathrm{tr}(A)<2$. Two cases arise.

Case (*a*): φ_1 and φ_2 are linearly dependent. In this case $P_1=P_2$ ($=P$ say), and if $A=P$, then $\mathrm{tr}(A)=1$.

Case (*b*): φ_1 and φ_2 span a two-dimensional subspace S of $\mathscr{H}$. It is obviously enough to consider the case when $S=\mathscr{H}$. Let φ_2' be a unit vector of S orthogonal to φ_1. Then for some $a, b \in \mathbf{D}$, $\varphi_2=a\varphi_1+b\varphi_2'$. Neither a nor b is zero and $|a|^2+|b|^2=1$. Write $\psi=|a|a^{-1}\varphi_2$ and

$$\psi_2 = |a|a^{-1}|b|^{-1}b\cdot\varphi_2'.$$

Then $\{\varphi_1,\psi_2\}$ is an orthonormal basis for S and $\psi=|a|\varphi_1+|b|\psi_2$. Clearly P_2 is the projection on the subspace $\mathbf{D}\cdot\psi$. With respect to the orthonormal basis $\{\varphi_1,\psi_2\}$, P_1 and P_2 have the matrices given by

$$P_1 \sim \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \qquad P_2 \sim \begin{pmatrix} |a|^2 & |a||b| \\ |a||b| & |b|^2 \end{pmatrix}.$$

Since $|a|^2+|b|^2=1$, the eigenvalues of the matrix of P_2-P_1 are $\pm|b|$, and hence there exists a unitary operator V such that the operator $V(P_2-P_1)V^{-1}$ has the matrix given by

$$V(P_2-P_1)V^{-1} \sim \begin{pmatrix} |b| & 0 \\ 0 & -|b| \end{pmatrix}.$$

Let M be the unique operator such that VMV^{-1} has the matrix given by

$$VMV^{-1} \sim \begin{pmatrix} |b| & 0 \\ 0 & 0 \end{pmatrix}.$$

Obviously M is self-adjiont, ≥ 0 and $\mathrm{tr}(M)=|b|<1$. Further,

$$VMV^{-1}-V(P_2-P_1)V^{-1} \geq 0$$

and hence $M \geq P_2-P_1$. Let $A=P_1+M$. Then $P_1 \leq A$, $P_2 \leq A$ but $\mathrm{tr}(A)=1+|b|<2$. This proves the lemma.

Theorem 4.33. *Let* $\mathscr{H}$ *be a separable infinite dimensional Hilbert space over* $\mathbf{D}$ *and* T *a symmetry of* $\mathscr{H}$*. For any* $U \in \mathscr{W}$ *let* p_U *be the state* $M \rightarrow \mathrm{tr}(P^M U)$ *of* $\mathscr{L}$ *and* $U_T=TUT^{-1}$*. If*

$$\alpha_T(p_U) = p_{U_T}, \tag{92}$$

then $\alpha_T(p_U \rightarrow p_{U_T})$ *is a convex automorphism of* $\mathscr{S}$*. Conversely, let* α *be a convex automorphism of* $\mathscr{S}$*. Then there exists a symmetry* T *of* $\mathscr{H}$ *such that*

$$\alpha = \alpha_T. \tag{93}$$

T *is determined up to multiplication by a number* $c \in \mathbf{D}$ *with* $|c|=1$.

Proof. The first part of the theorem is essentially the content of lemma 4.30. To prove the converse, let α be a convex automorphism of $\mathscr{S}$. By theorem 4.23, α induces a convex automorphism of $\mathscr{W}$, denoted once again by α. Let $\alpha^{\sim}$ be the map of $\mathscr{W}_+$, determined by α (cf. lemma 4.31). Since the extreme points of $\mathscr{W}$ are the orthogonal projections on the one-dimensional subspaces of $\mathscr{H}$ and since α is convex, α induces a one-one correspondence of the set $\mathscr{R}$, of all rays of $\mathscr{H}$, onto itself. Let $\alpha(\mathbf{r} \rightarrow \mathbf{r}^{\alpha})$ be this correspondence. Since $\alpha^{\sim}$ preserves the trace and the partial ordering $\leq$ of $\mathscr{W}_+$ (by (iv) of lemma 4.31), and since the orthogonality of two rays has been completely characterized in terms of the trace function and the partial ordering $\leq$ in lemma 4.32, it follows at once that $\mathbf{r} \to \mathbf{r}^\alpha$ is orthogonality-preserving. By theorem 4.29 there exists a symmetry T such that $\mathbf{r}^\alpha = T[\mathbf{r}]$ for all rays $\mathbf{r}$. If α_T is the convex automorphism of $\mathscr{S}$ associated with T, this means that α_T and α coincide on the set of extreme points of $\mathscr{W}$. From this it follows that $\alpha = \alpha_T$. The fact that T is determined up to multiplication by an element $c \in \mathbf{D}$ with $|c| = 1$ now follows from theorem 4.29. The theorem is completely proved.

4. LOGICS ASSOCIATED WITH VON NEUMANN ALGEBRAS

If the logic $\mathscr{L}$ of a quantum mechanical system is not isomorphic to a standard one, then the question of describing its observables, states, and symmetries becomes a more difficult one. There are no general results in this connection. It is our aim to give a few examples which illustrate the wide possibilities as well as the difficulties involved in the construction of any general theory.

A large class of logics which are sublogics of standard logics may be constructed using the theory of von Neumann algebras. We confine ourselves to the case of *complex scalars*. Let $\mathbf{D} = \mathbf{C}$ be the complex number field and $\mathscr{H}$ a separable infinite dimensional Hilbert space over $\mathbf{C}$. We write $\mathfrak{A}(\mathscr{H})$ for the algebra of all operators on $\mathscr{H}$. For $u, v \in \mathscr{H}$, let

$$\lambda_{u,v} : A \to \langle Au, v \rangle \qquad (A \in \mathfrak{A}(\mathscr{H})).$$

Then the $\lambda_{u,v}$ are linear functionals on $\mathfrak{A}(\mathscr{H})$. The *weak topology* on $\mathfrak{A}(\mathscr{H})$ is the smallest one with respect to which all the $\lambda_{u,v}$ are continuous. A *von Neumann algebra* $\mathfrak{A}$ is a subalgebra (containing I) of $\mathfrak{A}(\mathscr{H})$ such that (i) if $A \in \mathfrak{A}$, then $A^* \in \mathfrak{A}$, and (ii) $\mathfrak{A}$ is a closed subset of $\mathfrak{A}(\mathscr{H})$ in the weak topology. If $\mathfrak{A}'$ is the set of all elements of $\mathfrak{A}(\mathscr{H})$ commuting with all the elements of $\mathfrak{A}$, $\mathfrak{A}'$ is also a von Neumann algebra and $(\mathfrak{A}')' = \mathfrak{A}$ (cf. Dixmier [1] for the general theory of von Neumann algebras).

Let $\mathfrak{A}$ be a von Neumann algebra. We define

$$\mathscr{L}_{\mathfrak{A}} = \{M : M \in \mathscr{L}(\mathscr{H}), P^M \in \mathfrak{A}\}. \tag{94}$$

It is then easy to verify that $\mathscr{L}_{\mathfrak{A}}$ is a logic. We shall call it the logic associated with $\mathfrak{A}$. A closed linear manifold M lies in $\mathscr{L}_{\mathfrak{A}}$ if and only if every element of $\mathfrak{A}'$ leaves M invariant. Thus the logic $\mathscr{L}_{\mathfrak{A}}$ may also be introduced as the logic of closed invariant linear manifolds of some von Neumann algebra.

The one-one correspondence $x \to A_x$ between observables of $\mathscr{L}_{\mathfrak{A}}$ and self-adjoint operators on $\mathscr{H}$, which was discussed in section 2 of this chapter, clearly persists even in this case. However, not every self-adjoint operator is the operator of an observable associated with $\mathscr{L}_{\mathfrak{A}}$.

Let A be a self-adjoint operator, not necessarily bounded. Then $A = A_x$ for an observable x of $\mathscr{L}_{\mathfrak{A}}$ if and only if the spectral projections of A lie in $\mathfrak{A}$. If A is bounded, the condition will then be that $A \in \mathfrak{A}$. If A is unbounded, the condition will have to be stated more carefully. In this case, $A = A_x$ for an $\mathscr{L}_{\mathfrak{A}}$-observable if and only if $UAU^{-1} = A$ for all unitary operators $U \in \mathfrak{A}'$ (cf. Dixmier [1]; such A's are said to be *affiliated* to $\mathfrak{A}$). $\mathscr{C} = \mathfrak{A} \cap \mathfrak{A}'$ is the center of the algebra $\mathfrak{A}$. If we write $\mathscr{L}_{\mathscr{C}} = \mathscr{C} \cap \mathscr{L}_{\mathfrak{A}}$, then, $\mathscr{L}_{\mathscr{C}}$ is the center of the logic $\mathscr{L}_{\mathfrak{A}}$. $\mathscr{L}_{\mathfrak{A}}$ is irreducible if and only if $\mathscr{L}_{\mathscr{C}} = \{0, I\}$, or, equivalently, if $\mathscr{C}$ consists only of multiples of the identity, i.e., in the standard terminology of von Neumann algebras, if and only if $\mathfrak{A}$ is a factor.

It was Murray and von Neumann [1] who pioneered the entire study of factors. They classified factors into three types, I, II, and III. A factor $\mathfrak{A}$ is of type I if and only if $\mathfrak{A}$ is isomorphic to the algebra $\mathfrak{A}(\mathscr{K})$ of some Hilbert space $\mathscr{K}$. From our point of view, $\mathfrak{A}$ is a factor of type I if and only if $\mathscr{L}_{\mathfrak{A}}$ is isomorphic to the standard logic $\mathscr{L}(\mathscr{K})$ for some $\mathscr{K}$. The results of sections 4.2 and 4.3 therefore give a complete description of the sets of observables and states when $\mathscr{L} = \mathscr{L}_{\mathfrak{A}}$ for some factor $\mathfrak{A}$ of type I. When $\mathfrak{A}$ is a factor but not of type I, no similar theory exists. As an example, we may mention that the problem of determining all the states of $\mathscr{L}_{\mathfrak{A}}$, when $\mathfrak{A}$ is a factor type II or III.

We have remarked previously that if $\mathfrak{A}$ is not a factor, the center $\mathscr{L}_{\mathscr{C}}$ of $\mathscr{L}_{\mathfrak{A}}$ is nontrivial. We have also seen in Chapter III that there is an intimate connection between $\mathscr{L}_{\mathscr{C}}$ and the pure states of $\mathscr{L}_{\mathfrak{A}}$. It is clear from theorem 3.19 that if any contact is to be made with conventional quantum theory, $\mathscr{L}_{\mathscr{C}}$ cannot be continuous. At the other extreme lies the situation in which $\mathscr{L}_{\mathscr{C}}$ is *discrete*, i.e., when there exist mutually orthogonal elements $M_1, M_2, \cdots$ in $\mathscr{L}_{\mathscr{C}}$ such that (i) $\bigvee_n M_n = \mathscr{H}$, and (ii) $M \in \mathscr{L}_{\mathscr{C}}$ if and only if for some $n_1 < n_2 < \cdots, M = \bigvee_k M_{n_k}$. The M_i's are then the minimal elements of $\mathscr{L}_{\mathscr{C}}$. A case when $\mathscr{L}_{\mathscr{C}}$ is discrete is obtained when we choose M_n as above, and define $\mathfrak{A}$ to be the set of all operators which leave each M_n invariant. Then, for a closed linear manifold M of $\mathscr{H}$, $M \in \mathscr{L}_{\mathfrak{A}}$ if and only if $M = \bigvee_n (M \wedge M_n)$. If we assume that $\dim M_n \geq 3$ for all n, we can describe all the pure states of $\mathscr{L}_{\mathfrak{A}}$. In fact, from theorem 3.19 and Gleason's theorem we conclude immediately that p is pure if and only if there exists an n and a unit vector φ in M_n such that

$$p(M) = \|P^M \varphi\|^2 \qquad (M \in \mathscr{L}_{\mathfrak{A}}). \tag{95}$$

In other words, the pure states of $\mathscr{L}_{\mathfrak{A}}$ are still in one-one correspondence with the rays of $\mathscr{H}$; however, not every ray can be used. *Only the rays which belong to some M_n correspond to pure states.* Physicists describe this situation by saying that no ray of $\mathscr{H}$, which is a nontrivial superposition of rays in *distinct* M_n, can describe a (pure) state of $\mathscr{L}_{\mathfrak{A}}$. Each M_n is called a *coherent* subspace of (pure) states. The conditions which demand that the rays describing the pure states should lie in M_n for some n are usually referred to as *superselection* rules. (cf. Wick, Wightman, and Wigner [1]).

5. ISOMORPHISM AND IMBEDDING THEOREMS

It is very important in the foundations of quantum mechanics to examine the extent of generality of the concept of an abstract logic. In particular, the question arises whether every logic can be regarded as a sublogic of a standard logic. Definitive answers to such questions are not known. We shall prove in this section some theorems which analyze the structure of a somewhat restricted class of logics. The main theorem of this section was proved by Piron [1]. A related but somewhat weaker result was obtained by Zierler [1].

The main tool of our analysis is the coordinatization theorem of Chapter II. In order to be able to apply that theorem we must impose a sharp restriction on the class of logics to be analyzed. We introduce accordingly the following definition. A logic $\mathscr{L}$ is said to be *projective* if the following conditions are satisfied:

(i) given $a \neq 0$ in $\mathscr{L}$, there is a point $x < a$;
(ii) if $a \neq 0$ in $\mathscr{L}$ is the (lattice) sum of a finite set of points, then $\mathscr{L}[0,a]$ is a geometry of finite rank; we shall say that a is a
(96) *finite* element of $\mathscr{L}$ and write $\dim(a)$ for the dimension of $\mathscr{L}[0,a]$;
(iii) if $x, a \in \mathscr{L}$, $a \neq 0$, $\neq 1$ and x is a point, then there are points $y, z \in \mathscr{L}$ such that $y < a$, $z < a^{\perp}$ and $x < y \vee z$;
(iv) there exists at least one $a \in \mathscr{L}$ such that $4 \leq \dim(a) < \infty$.

For any $a \in \mathscr{L}$, let

$$\mathfrak{P}(a) = \{x : x \text{ a point}, x < a\}. \tag{97}$$

Note that if $\mathscr{L}$ is projective and its lattice is complete, then every element a of $\mathscr{L}$ is the lattice sum of the points it contains. In fact, if b is the sum of these points and $b \neq a$, $c = b^{\perp} \wedge a \neq 0$ and there would exist a point $x < c$ by (i) of (96) which contradicts the definition of b.

Lemma 4.34. *Let $\mathscr{L}$ be a projective logic whose underlying lattice is complete. If a and b_j $(j \in J)$ are elements of $\mathscr{L}$ and if $a \perp b_j$ for all j, $a \perp \bigvee_j b_j$. We have the usual identities:*

$$\Big(\bigvee_j b_j\Big)^{\perp} = \bigwedge_j b_j^{\perp},$$
$$\Big(\bigwedge_j b_j\Big)^{\perp} = \bigvee_j b_j^{\perp}.$$

If $a, b \in \mathscr{L}$, *then* $a < b$ *if and only if* $\mathfrak{P}(a) \subseteq \mathfrak{P}(b)$; $a \perp b$ *if and only if* $\mathfrak{P}(a) \perp \mathfrak{P}(b)$.

Proof. The first two assertions are proved exactly like the corresponding assertions in lemma 3.1. We therefore come to the proof of the last assertion. If $a < b$, $\mathfrak{P}(a) \subseteq \mathfrak{P}(b)$ obviously. On the other hand, let $a \not< b$. Then $a \wedge b \neq a$ and hence, by (96) (i), there exists a point $x \perp a \wedge b$ such that $x \in \mathfrak{P}(a)$. Clearly $x \notin \mathfrak{P}(b)$. If $a \perp b$, $\mathfrak{P}(a) \perp \mathfrak{P}(b)$ obviously. Suppose now that $\mathfrak{P}(a) \perp \mathfrak{P}(b)$. Then for any $x \in \mathfrak{P}(a)$, $x \perp \bigvee_{y \in \mathfrak{P}(b)} y = b$. Therefore $b \perp \bigvee_{x \in \mathfrak{P}(a)} x = a$.

Let $\mathbf{D}$ be a division ring and V a vector space, not necessarily finite dimensional, over $\mathbf{D}$ with dimension at least 4. Let θ be an involutive anti-automorphism of $\mathbf{D}$ and let $\langle . , . \rangle$ be a θ-bilinear symmetric form on $V \times V$. Let us assume that $\langle . , . \rangle$ is *definite*; i.e., $\langle x,x \rangle = 0$ if and only if $x = 0$. For any set M of vectors we write

$$M^{\perp} = \{u : \langle x,u \rangle = 0 \quad \text{for all} \quad x \in M\}. \tag{98}$$

Obviously, $M^{\perp}$ is a linear manifold in V, $M \cap M^{\perp} = 0$, and $M \subseteq M^{\perp\perp}$. If M is a linear manifold and $M = M^{\perp\perp}$, we shall say that M is $\langle . , . \rangle$-closed. *Note that no topology is involved in this definition.* The pair $(V, \langle . , . \rangle)$ is said to be *Hilbertian* if for any closed linear manifold M, one has:

$$V = M + M^{\perp} \tag{99}$$

algebraically. For any set $M \subseteq V$ we write

$$M^{-} = M^{\perp\perp} \tag{100}$$

If $A, B \subseteq V$ we write $A \perp B$ whenever $\langle a,b \rangle = 0$ for all $a \in A$ and $b \in B$. Clearly $0^{-} = 0$, $V^{-} = V$, and $0^{\perp} = V$, $V^{\perp} = 0$. If M and N are linear manifolds and $M \subseteq N$, then $N^{\perp} \subseteq M^{\perp}$. If $\{M_j\}$ is a collection of linear manifolds, M their algebraic sum, and $N \perp M_j$ for all j, then $N \perp M$. All this is trivial.

Lemma 4.35. *The following statements on a set* $M \subseteq V$ *are equivalent.*

(i) M is a $\langle . , . \rangle$*-closed linear manifold.*
(ii) $M = M^{-}$.
(iii) $M = N^{\perp}$ *for some set* $N \subseteq V$.

Moreover, for any set $M \subseteq V$,

$$M^{\perp\perp\perp} = M^{\perp}, \tag{101}$$

$$(M^{-})^{\perp} = M^{\perp}, \tag{102}$$

and

$$(M^{-})^{-} = M^{-}. \tag{103}$$

Finally, if $\{M_j\}$ *is a collection of* $\langle . , . \rangle$*-closed linear manifolds,* $\bigcap_j M_j$ *is also a* $\langle . , . \rangle$*-closed linear manifold.*

Proof. Clearly for any set $M \subseteq V$, $M \subseteq M^{\perp\perp}$. Thus $M^{\perp} \subseteq (M^{\perp})^{\perp\perp}$. On the other hand, as $M \subseteq M^{\perp\perp}$, $(M^{\perp\perp})^{\perp} \subseteq M^{\perp}$. Therefore $M^{\perp} = M^{\perp\perp\perp}$, which is (101). Since $M^{-} = M^{\perp\perp}$, (102) follows from (101). Further,

$$(M^{-})^{-} = (M^{\perp\perp})^{\perp\perp} = (M^{\perp\perp\perp})^{\perp} = M^{\perp\perp}$$

by (101), so that $(M^{-})^{-} = M^{-}$. This proves (103). We now come to the proof of the equivalence of (i) through (iii). Suppose M is a $\langle . , . \rangle$-closed linear manifold. Then $M = M^{\perp\perp}$, so that $M = M^{-}$. Thus (i) ⇒ (ii). If $M = M^{-}$, then $M = M^{\perp\perp} = (M^{\perp})^{\perp}$, proving that (ii) ⇒ (iii). Suppose now that $M = N^{\perp}$. Then, by (98), $M^{\perp\perp} = N^{\perp\perp\perp} = N^{\perp} = M$, proving (iii) ⇒ (i). Finally, let $M = \bigcap_j M_j$. As $M \subseteq M_j$, $M^{\perp\perp} \subseteq M_j^{\perp\perp} = M_j$. Thus

$$M^{\perp\perp} \subseteq \bigcap_j M_j = M.$$

This proves that M is a $\langle . , . \rangle$-closed linear manifold.

Corollary 4.36. *If* $M \subseteq V$, M^{-} *is the smallest* $\langle . , . \rangle$*-closed linear manifold containing* M*. In particular, if* $M \subseteq N$*, then* $M^{-} \subseteq N^{-}$.

Proof. M^{-} is $\langle . , . \rangle$-closed by (103). It is therefore enough to prove that if N is a $\langle . , . \rangle$-closed linear manifold and $M \subseteq N$, then $M^{-} \subseteq N$. But $M^{-} = M^{\perp\perp} \subseteq N^{\perp\perp} = N$. Finally if $M \subseteq N$, $M \subseteq N^{-}$, and as N^{-} is $\langle . , . \rangle$-closed, $M^{-} \subseteq N^{-}$.

Lemma 4.37. *If* M *is a finite dimensional subspace of* V*, then* M *is* $\langle . , . \rangle$*-closed.*

Proof. It is enough to prove that $M^{\perp\perp} \subseteq M$. Let $x \in M^{\perp\perp}$. Let N be the subspace spanned by M and x. $N \subseteq M^{\perp\perp}$. Since $\langle . , . \rangle$- is definite, its restriction to $N \times N$ is definite and hence nonsingular. But N is finite dimensional and hence it follows from (13) of Chapter II and the definiteness of $\langle . , . \rangle$ that N is the direct sum of M and $(M^{\perp} \cap N)$. Therefore, $x = x' + x''$ where $x' \in M$ and $x'' \in M^{\perp} \cap N$. Since $N \subseteq M^{\perp\perp}$, $x'' \in M^{\perp\perp} \cap M^{\perp}$ and hence $x'' = 0$. Thus $x \in M$.

Lemma 4.38. *If* $\{M_j\}$ *is a family of linear manifolds of* V*, and if* $\sum_j M_j$ *denotes the (algebraic) linear span of the* M_j,

$$\Big(\sum_j M_j\Big)^{-} = \Big(\bigcap_j M_j^{\perp}\Big)^{\perp}. \tag{104}$$

Proof. Let $M = \sum_j M_j$. Since $M_j \subseteq M$, $M^{\perp} \subseteq M_j^{\perp}$ so that $M^{\perp} \subseteq \bigcap_j M_j^{\perp}$. This shows that $M^{-} = M^{\perp\perp} \supseteq (\bigcap_j M_j^{\perp})^{\perp} = N$, say. On the other hand, as $\bigcap_j M_j^{\perp} \subseteq M_j^{\perp}$, $M_j \subseteq M_j^{-} \subseteq N$ for all j so that $M \subseteq N$. Since N is $\langle . , . \rangle$-closed, it follows from corollary 4.36 that $M^{-} \subseteq N$. Therefore $N = M^{-}$.

Lemma 4.39. *If* $(V, \langle . , . \rangle)$ *is Hilbertian and* $M_1, M_2, \cdots, M_s$ *is a finite set of mutually orthogonal* $\langle . , . \rangle$*-closed linear manifolds, then* $\sum_{j=1}^{s} M_j$ *is also a* $\langle . , . \rangle$*-closed linear manifold.*

Proof. It is enough to consider the case $s=2$ since the general case follows at once by induction. We want to prove that

$$(M_1+M_2)^- = M_1+M_2.$$

Write $M=M_1+M_2$. Let $x \in V$. Since $(V, \langle . , . \rangle)$ is Hilbertian, we have the unique decompositions

$$\begin{aligned} x &= x_1+y \qquad (x_1 \in M_1, y \in M_1{}^\perp), \\ y &= x_2+z \qquad (x_2 \in M_2, z \in M_2{}^\perp). \end{aligned}$$

Now, as $M_2 \subseteq M_1{}^\perp$, $x_2 \in M_1{}^\perp$ so that $z \in M_1{}^\perp$ also. Therefore,

$$z \in M_1{}^\perp \cap M_2{}^\perp,$$

and we have:

$$x = x_1+x_2+z \qquad (x_1 \in M_1, x_2 \in M_2, z \in M_1{}^\perp \cap M_2{}^\perp).$$

Suppose now that $x \in (M_1+M_2)^-$. Then the above equation shows that $z \in (M_1+M_2)^- \cap M_1{}^\perp \cap M_2{}^\perp$, which implies that $z=0$ on using (104). Therefore $(M_1+M_2)^- = M_1+M_2$.

Theorem 4.40. *Let* **D** *be a division ring,* **V** *a vector space over* **D** *with* $4 \leq \dim V < \infty$, θ *an involutive anti-automorphism of* **D**, *and* $\langle . , . \rangle$ *a definite, symmetric,* θ*-bilinear form on* $V \times V$. *Let* $\mathscr{L}(V, \langle . , . \rangle)$ *be the set of* $\langle . , . \rangle$*-closed linear manifolds of* V, *partially ordered under inclusion. If* $(V, \langle . , . \rangle)$ *is Hilbertian, then* $\mathscr{L}(V, \langle . , . \rangle)$ *is a complete projective logic, and for any collection* $\{M_j\}$ *of* $\langle . , . \rangle$*-closed linear manifolds of* V, *the lattice operations in* $\mathscr{L}(V, \langle . , . \rangle)$ *are given by*

$$\begin{aligned} \bigvee_j M_j &= \Big(\sum_j M_j\Big)^-, \\ \bigwedge_j M_j &= \bigcap_j M_j. \end{aligned} \tag{105}$$

Conversely, let $\mathscr{L}$ *be any complete projective logic. Then there exists a division ring* **D**, *an involutive anti-automorphism* θ *of* **D**, *a vector space* V *over* **D**, *and a definite symmetric* θ*-bilinear form* $\langle . , . \rangle$ *on* $V \times V$ *such that* $(V, \langle . , . \rangle)$ *is Hilbertian and* $\mathscr{L}$ *is isomorphic to* $\mathscr{L}(V, \langle . , . \rangle)$.

Proof. Let **D**, θ, V, $\langle . , . \rangle$ be as given in the first half of the theorem. Let $\mathscr{L} = \mathscr{L}(V, \langle . , . \rangle)$ be the partially ordered set of $\langle . , . \rangle$-closed linear manifolds of V. Assume that $(V, \langle . , . \rangle)$ is Hilbertian. For any linear manifold M, M^- is the smallest element of $\mathscr{L}$ containing it. It is clear from lemmas 4.35 through 4.38 that $\mathscr{L}$ is a complete lattice with the lattice operations given by (105). Next we observe that if we define, for any $M \in \mathscr{L}$, $M^\perp$ by (97), then $\mathscr{L}$ becomes a logic. To check this, the only thing not immediately obvious is the weak modularity (2) (ii) of Chapter III. Suppose then $M_1, M_2 \in \mathscr{L}$ and $M_1 \subseteq M_2$. Let $N = M_1^\perp \cap M_2$. By lemma 4.35, N is a $\langle . , . \rangle$-closed linear manifold so that $N \in \mathscr{L}$. Clearly $N \subseteq M_2$, $N \perp M_1$. Since $(V, \langle . , . \rangle)$ is Hilbertian, $V = M_1 + M_1^\perp$ and for any $x \in V$ we have the decomposition $x = x_1 + x_2$, where $x_1 \in M_1$, $x_2 \in M_1^\perp$. If now $x \in M_2$, then x_1 and $x_2 \in M_2$ also, so that $x_2 \in N$. Therefore $M_2 = M_1 + N$ and *a fortiori* $M_2 = M_1 \vee N$. $\mathscr{L}$ is thus a complete logic. We claim that it is projective. By lemma 4.37, $\mathscr{L}$ contains all finite dimensional subspaces. As dim $V \geq 4$, (i), (ii), and (iv) of (96) follow at once from this fact; (iii) follows from the *algebraic* decomposition $V = M + M^\perp$ for any $M \in \mathscr{L}$. $\mathscr{L}$ is thus a complete projective logic.

We now come to the converse. Let $\mathscr{L}$ be a complete projective logic. Let us now define $\mathscr{L}'$ by

$$(106) \qquad \mathscr{L}' = \{a : a \in \mathscr{L}, \quad a \text{ finite}\}.$$

By (ii) of (96), $a \in \mathscr{L}'$ if and only if the underlying lattice of $\mathscr{L}[0,a]$ is a geometry. Under the induced partial ordering, $\mathscr{L}'$ is thus a *generalized geometry*. Therefore there exists by theorem 2.16, a division ring **D**, a vector space V of dimension at least 4 over **D**, and an isomorphism of $\mathscr{L}'$ onto the generalized geometry of all finite dimensional subspaces of V. Let γ denote any such isomorphism.

Fix a nonzero vector $v_0 \in V$. Let W be any finite dimensional subspace of V containing v_0 and of dimension at least 3. The isomorphism γ transfers the orthocomplementation in $\mathscr{L}[0,a]$ into an orthocomplementation on the projective geometry of subspaces of W. Hence, by theorem 2.7 there exists an involutive anti-automorphism θ_W of **D** and a symmetric, definite θ_W-bilinear form $\langle . , . \rangle_W$ on $W \times W$, inducing the orthocomplementation in question, with $\langle v_0, v_0 \rangle_W = 1$. We now argue as in lemma 4.2 to conclude that $\theta_W = \theta$ is independent of W and that there exists a symmetric, definite θ-bilinear form $\langle . , . \rangle$ on $V \times V$ such that (i) $\langle v_0, v_0 \rangle = 1$, and (ii) if x and y are points of $\mathscr{L}$, $x \perp y$ if and only if the rays $\gamma(x)$ and $\gamma(y)$ are orthogonal with respect to $\langle . , . \rangle$.

Let $\mathscr{L}^\sim = \mathscr{L}(V, \langle . , . \rangle)$ be the lattice of all $\langle . , . \rangle$-closed linear manifolds of V. We shall now extend γ to an isomorphism $\gamma^\sim$ of $\mathscr{L}$ onto $\mathscr{L}^\sim$. For any $a \in \mathscr{L}$, let

$$(107) \quad \gamma^\sim(a) = \{v : v \in V \quad \text{and} \quad v \in \gamma(x) \quad \text{for some point } x < a\}.$$

We shall first show that $\gamma^\sim(a)$ is a linear manifold in V. This is obvious for $a = 0$ or $a = 1$, as then $\gamma^\sim(0) = 0$ and $\gamma^\sim(1) = V$. Let then $a \neq 0$, $\neq 1$. If $v_1, v_2 \in \gamma^\sim(a)$, there are points $x_1, x_2 < a$ such that $v_j \in \gamma(x_j)$, $j = 1, 2$. Now, as γ is an isomorphism of $\mathscr{L}'$ onto the generalized geometry of finite dimensional subspaces of V, $\gamma(x_1)$ and $\gamma(x_2)$ are rays in $\gamma(x_1 \vee x_2)$. Thus if $v = c_1 v_1 + c_2 v_2$ where $c_1, c_2 \in \mathbf{D}$, v lies in $\gamma(x_1 \vee x_2)$ and hence $v \in \gamma(x)$ for some point $x < x_1 \vee x_2 < a \cdot \gamma^{\sim}(a)$ is thus a linear manifold. We shall prove next that $\gamma^{\sim}(a)$ is $\langle . , . \rangle$-closed. From lemma 4.34 it follows easily that for a point $x \in \mathscr{L}$, $x < a$ if and only if $x \perp y$ for all $y \in \mathfrak{P}(a^{\perp})$. Therefore,

$$(108) \qquad \gamma^{\sim}(a) = (\gamma^{\sim}(a^{\perp}))^{\perp},$$

which shows that $\gamma^{\sim}(a)$ is $\langle . , . \rangle$-closed. Moreover, as (108) is valid for all a (in particular, for $a^{\perp}$), we have:

$$(109) \qquad \gamma^{\sim}(a^{\perp}) = (\gamma^{\sim}(a))^{\perp}.$$

(109) shows that $\gamma^{\sim}$ preserves orthocomplementation. From (107) we conclude, on using lemma 4.34, that for $a, b \in \mathscr{L}$, $\gamma^{\sim}(a) \subseteq \gamma^{\sim}(b)$ if and only if $a < b$. In particular, $\gamma^{\sim}(a) = \gamma^{\sim}(b)$ if and only if $a = b$. Therefore, in order to prove that $\gamma^{\sim}$ is a lattice isomorphism of $\mathscr{L}$ onto $\mathscr{L}(V, \langle . , . \rangle)$, it remains only to prove that $\gamma^{\sim}$ is surjective. Let M be any $\langle . , . \rangle$-closed linear manifold of V and let the subsets A and B of $\mathscr{L}$ be defined as follows:

$$A = \{x : x \in \mathscr{L},\ x \text{ is a point such that } \gamma(x) \subseteq M\},$$
$$B = \{x : x \in \mathscr{L},\ x \text{ is a point such that } \gamma(x) \subseteq M^{\perp}\}.$$

Since $M = M^{\perp\perp}$, $x \in A$ if and only if $x \perp B$. Let a and b be the lattice sums of the points of A and B, respectively; these exist as $\mathscr{L}$ is complete. It follows from lemma 4.34 rather easily that a and b are orthogonal complements of each other and $A = \mathfrak{P}(a)$, $B = \mathfrak{P}(b)$. The equation $A = \mathfrak{P}(a)$ implies that $\gamma^{\sim}(a) = M$. The proof that $\gamma^{\sim}$ is surjective is complete. $\gamma^{\sim}$ is thus a lattice isomorphism of $\mathscr{L}$ onto $\mathscr{L}(V, \langle . , . \rangle)$.

We observe finally that $(V, \langle . , . \rangle$ is Hilbertian. In fact, let M, $\neq 0$, $\neq V$, be any $\langle . , . \rangle$-closed linear manifold, and $v \in V$, $v \neq 0$. Let $a \in \mathscr{L}$ be such that $\gamma^{\sim}(a) = M$. Clearly, $a \neq 0$, $\neq 1$. Then $\gamma^{\sim}(a^{\perp}) = M^{\perp}$. Choose $x \in \mathscr{L}$ such that $v \in \gamma(x)$. By (iii) of (96), there are points y, z of $\mathscr{L}$ such that $y < a$, $z < a^{\perp}$ and $x < y \vee z$. Hence $\gamma(x) < \gamma(y) \vee \gamma(z) = \gamma(y) + \gamma(z)$. This shows that $v \in \gamma(y) + \gamma(z) \subseteq M + M^{\perp}$ and demonstrates the Hilbertian character of $(V, \langle . , . \rangle)$. This completes the proof of the theorem.

The division ring $\mathbf{D}$ which enters the second half of theorem 4.40 is uniquely determined by $\mathscr{L}$ (up to isomorphism of course), as can be easily deduced from theorem 2.1. If $\mathbf{D}$ were one of $\mathbf{R}$, $\mathbf{C}$, or $\mathbf{H}$, and $\mathscr{L}$ is a complete projective logic, we shall say that $\mathscr{L}$ is *associated with* $\mathbf{D}$ if the involutive anti-automorphism θ, constructed in theorem 4.40, is continuous. This is a restriction only in the complex case. If $\mathscr{L}$ is associated with $\mathbf{D}$, θ must coincide with the conjugation $*$, which is complex conjugation when $\mathbf{D} = \mathbf{C}$, the identity when $\mathbf{D} = \mathbf{R}$, and the standard conjugation when $\mathbf{D} = \mathbf{H}$. In view of theorem 4.40 and these comments we can assert that, given any complete projective logic $\mathscr{L}$ associated with $\mathbf{D}$ which is one of $\mathbf{R}$, $\mathbf{C}$, or $\mathbf{H}$, there exists a pre-Hilbert space V, with inner product $\langle . , . \rangle$, such that (i) $(V, \langle . , . \rangle)$ is Hilbertian and (ii) $\mathscr{L}$ is isomorphic to the complete projective logic of all $\langle . , . \rangle$-closed linear manifolds of V. It is natural to expect that V is then actually a Hilbert space. This essentially leads to Piron's theorem. We shall obtain it as a consequence of a few lemmas.

It is an interesting problem to examine what natural assumptions on a geometry imply that the associated division ring is isomorphic to one of $\mathbf{R}$, $\mathbf{C}$, $\mathbf{H}$. This is a classical question and is intimately connected with the topologies on a geometry (cf. Kolmogorov [1], Weiss and Zierler [1]).

Lemma 4.41. *Let V be a pre-Hilbert space over* $\mathbf{D}$ (*one of* $\mathbf{R}$, $\mathbf{C}$, *or* $\mathbf{H}$) *and let $\mathscr{H}$ be its completion. If $x_0, x_0' \in \mathscr{H}$ are orthogonal, there exists a pair of sequences x_n and x_n' such that*

$$\text{(110)} \qquad \begin{array}{ll} \text{(i) } x_n \perp x_m' & (n, m = 0, 1, 2, \cdots), \\ \text{(ii) } x_n, x_m' \in V & \text{and} \quad x_n \to x_0, x_m' \to x_0'. \end{array}$$

Proof. We shall prove first that if $z \in \mathscr{H}$ is a nonzero vector, then $\{z\}^\perp \cap V$ is dense in $\{z\}^\perp$. In fact, let $y \perp z$, and let $\{y_n\}$, $\{z_n\}$ be sequences of elements of V such that $y_n \to y$, $z_n \to z$. Write

$$\text{(111)} \qquad y_n' = y_n - \langle y_n, z\rangle\langle z_n, z\rangle^{-1} z_n.$$

Then $y_n' \in V$ and $\langle y_n', z\rangle = 0$ for all n. As $y_n \to y$ and $z_n \to z$,

$$\langle y_n, z\rangle\langle z_n, z\rangle^{-1} z_n \to \langle y, z\rangle\langle z, z\rangle^{-1} z = 0,$$

showing that $y_n' \to y$. This proves our assertion. Suppose next that $y, y_1, \cdots, y_s$ are mutually orthogonal nonzero vectors. Applying repeatedly the result proved just now, we conclude in succession that $\{y_s\}^\perp \cap V$ is dense in $\{y_s\}^\perp$, $\{y_s, y_{s-1}\}^\perp \cap V$ is dense in $\{y_s, y_{s-1}\}^\perp$ and so on. In other words, given any finite dimensional subspace Y of $\mathscr{H}$ and a vector $y \perp Y$, there exists a sequence $\{z_n\}$ in V such that $z_n \perp Y$ for all n and $z_n \to y$.

This said, we come to the proof of the lemma. Let $x_0, x_0' \in \mathscr{H}$ be given with $x_0 \perp x_0'$. We shall show by induction that there are sequences $\{x_n\}$ and $\{x_n'\}$ in V such that

$$\text{(112)} \qquad \begin{array}{c} \|x_n - x_0\| \leq 2^{-n}, \qquad \|x_n' - x_0'\| \leq 2^{-n}, \\ x_n \perp x_m' \qquad (n, m = 0, 1, 2, \cdots). \end{array}$$

Suppose $x_1, x_2, \cdots, x_k$ and $x_1', x_2', \cdots, x_k'$ have been constructed so that (112) is satisfied for all $n, m = 0, 1, 2, \cdots, k$. By the result proved in the previous paragraph, there is a vector $x'_{k+1} \in V$ such that

$$x'_{k+1} \perp x_n \qquad (n = 0, 1, \cdots, k),$$
$$\|x'_{k+1} - x_0'\| \leq 2^{-(k+1)}.$$

Applying the same argument again, we can find $x_{k+1} \in V$ such that

$$x_{k+1} \perp x_n' \qquad (n = 0, 1, \cdots, k+1),$$
$$\|x_{k+1}-x_0\| \leq 2^{-(k+1)}.$$

Thus $\{x_n\}_{0 \leq n \leq k+1}$ and $\{x_n'\}_{0 \leq n \leq k+1}$ satisfy (112) for n and $m \leq k+1$. This proves by induction the existence of a pair of sequences satisfying (112). This immediately proves lemma 7.41.

Lemma 4.42.† *Let V be a pre-Hilbert space over* $\mathbf{D}$ *such that* $(V, \langle . , . \rangle)$ *is Hilbertian. Then V is complete.*

Proof. Let $\mathscr{H}$ be the completion of V and x_0 an arbitrary unit vector of $\mathscr{H}$. We denote, for any subset L of $\mathscr{H}$, its orthogonal complement in $\mathscr{H}$ by L'. We must prove that $x_0 \in V$. Since V is dense in $\mathscr{H}$, there is a $y \in V$ such that $\langle y,x_0 \rangle \neq 0$, and by changing y to a suitable multiple of it we may assume that $\langle y,x_0 \rangle = 1$. Let $z_0 = y - x_0$. Then $y = x_0 + z_0$ and $x_0 \perp z_0$. By lemma 4.41 there are sequences $\{x_n\}$ and $\{z_n\}$ in V such that

$$(113) \qquad x_n \to x_0, \quad z_n \to z_0, \quad x_n \perp z_m \qquad (n, m = 0, 1, 2, \cdots).$$

Let $M \subseteq V$ be the lattice sum in $\mathscr{L}(V, \langle . , . \rangle)$ of the rays $\mathbf{D} \cdot x_n$ $(n = 1, 2, \cdots)$. Then, as $(V, \langle . , . \rangle)$ is Hilbertian,

$$V = M + M^{\perp}.$$

Now $z_m \perp x_n$ for all n so that $z_m \in M^{\perp}$ for all m. Hence $x_0 \in M^{\mathrm{cl}}$ and $z_0 \in (M^{\perp})^{\mathrm{cl}}$ where cl denotes closure in $\mathscr{H}$. Moreover, it is obvious that $(M^{\perp})^{\mathrm{cl}} \subseteq (M^{\mathrm{cl}})'$. Therefore,

$$(114) \qquad y = x_0 + z_0, \quad x_0 \in M^{\mathrm{cl}}, \quad z_0 \in (M^{\mathrm{cl}})'.$$

On the other hand, as $V = M + M^{\perp}$, we can also write

$$(115) \qquad y = x_0' + z_0', \quad x_0' \in M, \quad z_0' \in M^{\perp}.$$

As $M^{\perp} \subseteq (M^{\mathrm{cl}})'$, a comparison of (114) and (115) shows that $x_0' = x_0$ and $z_0' = z_0$. In particular, $x_0 \in V$. This completes the proof of the lemma.

Lemma 4.43. *Let $\mathscr{L}$ be a projective logic with the property that any family of mutually orthogonal points of $\mathscr{L}$ is at most denumerable. Then $\mathscr{L}$ is complete.*

Proof. It is enough to prove that arbitrary lattice unions exist in $\mathscr{L}$. Let $\{a_j : j \in F\}$ be a family of elements of $\mathscr{L}$. For any countable subset $D \subseteq F$, let

$$c(D) = \bigvee_{j \in D} a_j.$$

---

† The argument for completeness is essentially that of I. Amemiya and H. Araki, Publ. Res. Inst. Math. Sci. Kyoto Univ. A2 (1966), p. 423. Their theorem is that if V is pre-Hilbertian and $\mathscr{L}(V, \langle . , . \rangle)$ is orthomodular, then V is complete.

Let B be a family of mutually orthogonal points x such that (i) for each $x \in B$ there is a countable set $D \subseteq F$ such that $x < c(D)$, and (ii) B is maximal with respect to (i). By hypothesis, B is at most countable. Let

$$B = \{x_1, x_2, \cdots\},$$

and for each n let $D_n \subseteq F$ be a countable set such that

$$x_n < c(D_n)$$

for all n. Write $D = \bigcup_n D_n$. If we define x by

$$x = \bigvee_n x_n,$$

then $x < c(D)$. If $x \neq c(D)$, there will exist, by virtue of (i) of (96), a point $y < c(D)$ such that $y \perp x$. Since D is countable, this contradicts the maximality of B. Therefore $x = c(D)$. We claim that $a_j < x$ for all $j \in F$. If $j \in D$, this is trivial. If $j \notin D$ and $a_j \nless x$, $a_j \vee x \neq x$, and hence by (i) of (96) there will exist a point t such that $t \perp x$ and $t < a_j \vee x$. Since $t < c(D \cup \{j\})$ and $t \perp B$, we have a contradiction. Consequently, a_j must be $< x$. This proves that $x = \bigvee_{j \in F} a_j$.

Theorem 4.44. (*Piron* [1]). *Let $\mathscr{L}$ be any logic. Then, a necessary and sufficient condition that $\mathscr{L}$ be isomorphic to the logic of all closed linear manifolds of a separable Hilbert space over* $\mathbf{D}$ (*which is one of* $\mathbf{R}$, $\mathbf{C}$, *or* $\mathbf{H}$) *is that $\mathscr{L}$ be a projective logic associated with* $\mathbf{D}$ *and have the property that every family of mutually orthogonal points of $\mathscr{L}$ be at most countable.*

Proof. The proof follows at once from theorem 4.40 and lemmas 4.41 through 4.43.

NOTES ON CHAPTER IV

1. *Hidden variables: Is the quantum mechanical description of microphysics complete?* The question of "hidden variables" has persisted in the foundations of Quantum Theory from its very birth. The basic assumption in Quantum Mechanics, namely that in any state of an atomic system a physical quantity can in general have no sharply defined value (even under idealized measuring conditions) but only a probability distribution of the possible values, appeared to many people to be in very sharp conflict with all classical experience, and hence unacceptable, unless it was interpreted as follows: one is not able to get a complete description of the physical state of the system; there are additional coordinates ("hidden variables") that cannot be measured; and therefore the statistical character of the results of the experiments is due to the averaging over these hidden variables. This view, which regards the Hilbert space description of quantum states as

fundamentally incomplete, is reinforced by the example of statistical mechanics where the situation is essentially of this character.

Von Neumann was among the earliest to analyze this question, and his conclusion was that *no mechanism of hidden variables could account for the statistical content of Quantum Mechanics as currently formulated.* That he regarded this question, as well as his resolution of it, to be of great importance is clear from the fact that he raises the issue already in the preface of his book (pp. ix, x), returns to it briefly (pp. 209–210) before devoting a substantial effort (pp. 295–328) to a detailed mathematical and physical treatment (all references are to the 1955 English translation by Robert T. Beyer of von Neumann's book). We shall briefly summarize his argument but suggest strongly that the reader study von Neumann's brilliantly presented analysis. Von Neumann's reasoning, roughly speaking, has two parts to it, one physical, the other mathematical; and they reinforce each other.

The physical side of his argumentation uses the technique of ensembles and begins by considering a statistical ensemble consisting of a large number of models of the physical system under consideration. If we reject the view of Quantum Mechanics, and suppose that the values of physical quantities in the ensemble have positive dispersions only because the ensemble has not been resolved in a sufficiently fine manner, we must admit that we should be able to carry out such a resolution into subensembles in which the dispersions are zero, or at least diminished from those of the original ensemble. The method of doing this is to measure the physical quantities in succession, replacing the ensemble each time to one of the subensembles where the quantity has a sharply defined value. But then, the Heisenberg uncertainty relations between complementary physical quantities operate in such a way that the precision achieved for a quantity in any stage is destroyed at the next stage when a complementary quantity is measured; *we thus do not get ahead.* However, one may take exception to this line of reasoning by pointing out that there could conceivably be other methods of penetrating to the dispersion free ensembles. To settle this point decisively one must therefore answer the following question: Given an ensemble in which certain physical quantities have positive dispersion, is it possible to resolve this ensemble into a superposition of two subensembles, different from the original one and distinct from each other? Ensembles which cannot be so resolved are the *pure* (*homogeneous* in von Neumann's terminology) ones, and the above question becomes the following: Is every pure ensemble dispersion free? If the answer is no, then the hidden variables interpretation must be given up.

To answer this question *in the conventional model of Quantum Mechanics* von Neumann begins with the very interesting observation that since an ensemble is characterized (statistically) by the expectation values of all the physical quantities in it, one may replace the ensemble in the mathematical argument by the corresponding expectation functional; we must remember that if A is any bounded observable, its probability distribution is completely determined by the knowledge of all the expected values

$$\operatorname{Exp}(A^n),\ n = 0, 1, 2, \ldots.$$

Thus, for von Neumann, an ensemble is the same as the corresponding real valued (expectation) functional

$$\operatorname{Exp}\colon B_0(\mathscr{H}) \to \mathbf{R}$$

on the space $B_0(\mathscr{H})$ of all bounded self-adjoint operators on the complex Hilbert space $\mathscr{H}$; then von Neumann assumes that it satisfies the conditions:

$$(1)\qquad \begin{aligned} &\text{(a)}\quad \operatorname{Exp}(1) = 1;\\ &\text{(b)}\quad \operatorname{Exp}(A^2) \geqslant 0, \quad \forall A \in B(\mathscr{H});\\ &\text{(c)}\quad \operatorname{Exp} \text{ is linear (over } \mathbf{R});\\ &\text{(d)}\quad \operatorname{Exp} \text{ is continuous in the strong operator topology.}\end{aligned}$$

Nowadays, one works with the space $B(\mathscr{H})$ of all bounded operators on $\mathscr{H}$; the conditions (a)–(c) define a state, (b) being equivalent to $\operatorname{Exp}(AA^\dagger) \geqslant 0$, $\forall A \in B(\mathscr{H})$; with (d), we have a *normal state*. He showed that the functionals Exp are in one–one correspondence with bounded operators U which are self-adjoint, $\geqslant 0$, and of trace 1, the correspondence being defined by

$$(2)\qquad \operatorname{Exp}(A) = \operatorname{tr}(UA).$$

The correspondence preserves convexity and so matches the pure states with the extreme points of the convex set of the U's. He verified that the latter are the one-dimensional projections, and showed by explicit calculation that these are *never* dispersion free. This concluded his proof.

One must emphasize that his reasoning explicitly assumes that *Quantum Mechanics is described by the conventional model.* It is possible to modify von Neumann's argument so as to allow for the presence of superselection rules, but the essence of the reasoning does not change. Many subsequent criticisms of von Neumann's proof are unjust because they do not take into account the above assumption in his proof, although von Neumann states it quite unmistakably (see the first sentence of the first paragraph on p. 210, and lines 17 + on p. 324; see also the remarks of L. van Hove, *Von Neumann's contributions to Quantum Theory*, pp. 95–99, in John von Neumann, 1903–1957, Bull. Amer. Math. Soc. **64** (1958), No. 3, Part 2). The mathematical content of von Neumann's argument may thus be summarized as follows: *$B(\mathscr{H})$ does not admit any dispersion free normal state.*

The first question that arises naturally now is whether von Neumann's definition of a state was unduly restrictive. Obviously the conditions (c), (d), of (1) need to be examined more closely. A conservative definition would replace (c) by

$$(3) \qquad (\mathrm{c}') \quad \mathrm{Exp}(A+B) = \mathrm{Exp}(A) + \mathrm{Exp}(B) \quad \text{when} \quad AB = BA.$$

Let us call maps $\lambda : B_0(\mathscr{H}) \to \mathbf{R}$ satisfying (a), (b) of (1) and (c′) of (3) *physical states*. It then follows that $|\lambda(A)| \leqslant \|A\|$ and that the restriction of λ to the projections defines a finitely additive probability measure on the logic $\mathscr{L}(\mathscr{H})$. Conversely, it follows from the spectral theorem that any finitely additive probability measure on $\mathscr{L}(\mathscr{H})$ is the restriction to the projections of a unique physical state of $B_0(\mathscr{H})$. In other words, physical states may be identified with finitely additive probability measures on $\mathscr{L}(\mathscr{H})$. Thus the strongest generalization (in this direction) of the von Neumann result would be to show that there are no finitely additive dispersion free probability measures on $\mathscr{L}(\mathscr{H})$. The case when $3 \leqslant \dim(\mathscr{H}) < \infty$ is immediate; for, by Gleason's theorem, every finitely additive probability measure on $\mathscr{L}(\mathscr{H})$ is the restriction to the projections of a functional $A \to \mathrm{tr}\,(UA)$ ($U \geqslant 0$, $\mathrm{tr}(U) = 1$), and hence von Neumann's result already applies here. If $\dim(\mathscr{H}) = \infty$, we use the isomorphism $\mathscr{H} \xrightarrow{\sim} \mathscr{H} \otimes \mathbf{C}^3$ to imbed $\mathscr{L}(\mathbf{C}^3)$ inside $\mathscr{L}(\mathscr{H})$. Since for a finitely additive probability measure, being dispersion free is equivalent to its being $\{0,1\}$-valued, this property persists on restriction from $\mathscr{L}(\mathscr{H})$ to $\mathscr{L}(\mathbf{C}^3)$, allowing us to use the previous argument to conclude that $\mathscr{L}(\mathscr{H})$ has no dispersion free physical states.

A few footnotes may be added to this brief discussion. In the first place, as we shall see later in these notes when we discuss Gleason's theorem and its generalizations, physical states on any von Neumann algebra without I_2 summands are actually states; in particular, this is true for $B(\mathscr{H})$. Secondly, there are actually states that are not normal. In fact, regarding $B(\mathscr{H})$ as a Banach space, the states form a compact (in the weak topology) convex subset of the dual of $B(\mathscr{H})$; if $(e_m)_{m \geqslant 1}$ is an orthonormal basis of $\mathscr{H}$, $\lambda_m(A) = (Ae_m, e_m)$ ($A \in B(\mathscr{H})$), and ν is a limit point of $\{\lambda_m\}_{m \geqslant 1}$ one can show that ν is a state which is not normal. Our final remark is to point out that it was Mackey who first realized that one should really prove von Neumann's result with physical states rather than states, or at least with countably additive probability measures on $\mathscr{L}(\mathscr{H})$; Gleason's work which showed that every countably additive probability measure on $\mathscr{L}(\mathscr{H})$ defines a normal state on $B(\mathscr{H})$ at least when $3 \leqslant \dim(\mathscr{H}) \leqslant \infty$, was inspired by Mackey's investigations. In addition to being a significant and aesthetically satisfying strengthening of the von Neumann result, this work and others that were inspired by it have led to a deeper understanding of the questions involved. The fact that physical states are the same as states was proved in full generality only recently, and the main motivation for reaching this result had always been the desire to generalize Gleason's theorem to more general von Neumann algebras than $B(\mathscr{H})$.

A second question that arises now is whether one can relax the assumption in the von Neumann–Mackey–Gleason result that the logic of the physical system is $\mathscr{L}(\mathscr{H})$. There is an extensive literature on this question and I shall limit myself to a brief discussion of two results, those of Neal Zierler and Michael Schlessinger (*Boolean imbeddings of orthomodular sets and Quantum Logic*, Duke Math. J., **32** (1965), pp. 251–262), and of Simon Kochen and E. P. Specker (*The problem of hidden variables in Quantum Mechanics*, J. Math. Mech., **17** (1967), pp. 331–348); for more detail and other references the reader should consult the lucid discussion of these problems in the book of Beltrametti and Cassinelli (of notes to Chapter III).

In both of these articles, the interpretation of what is meant by hidden variables is such that it implies the existence of some sort of imbedding of the proposition system into a Boolean algebra. So the mathematical issue in both cases comes down to showing that under suitable assumptions on the proposition systems such imbeddings do not exist.

In Z–S it is assumed that the set $\mathscr{L}$ of experimentally verifiable quantum propositions is partially ordered, has 0 and 1, and an orthocomplementation $\perp : \mathscr{L} \to \mathscr{L}$ ($0^{\perp} = 1, 1^{\perp} = 0, a \leqslant b \Rightarrow b^{\perp} \leqslant a^{\perp}, a^{\perp\perp} = a, a \vee a^{\perp} = 1, a \wedge a^{\perp} = 0$) which is weakly modular ($a \leqslant b \Rightarrow a^{\perp} \wedge b$ exists and $b = a \vee (a^{\perp} \wedge b)$); it is not assumed that $\mathscr{L}$ is a lattice. As usual if $a \leqslant b^{\perp}$ we say a and b are orthogonal; sums of mutually orthogonal elements exist in $\mathscr{L}$. If $x, y \in \mathscr{L}$ they are said to *commute* if we can write $x = x_1 \vee z$, $y = y_1 \vee z$ where x_1, y_1, and z are mutually orthogonal; this is possible if and only if x and y are contained in a Boolean subalgebra of $\mathscr{L}$ (recall that $B \subset \mathscr{L}$ is a Boolean subalgebra if $0, 1 \in B$; if $c, a, b \in B \Rightarrow c^{\perp} \in B$, $a \vee b$, $a \wedge b$ exist in $\mathscr{L}$ and belong to B; and if B becomes a Boolean algebra with respect to $\vee$, $\wedge$, $\perp$). If B is a Boolean algebra, an *imbedding* of $\mathscr{L}$ into B is a map h ($\mathscr{L} \to B$) such that

$$(4) \qquad \begin{array}{ll} \text{(a)} & h(0) = 0, \quad h(1) = 1; \\ \text{(b)} & x \leqslant y \Leftrightarrow h(x) \leqslant h(y); \\ \text{(c)} & h(x)^{\perp} = h(x^{\perp}); \\ \text{(d)} & h(x \vee y) = h(x) \vee h(y); \\ & h(x \wedge y) = h(x) \wedge h(y) \quad \text{if } x \text{ and } y \text{ commute.} \end{array}$$

Actually it is enough to require (b), (c) and the single set of relations $h(x+y) = h(x) \vee h(y)$ in (d). Z–S prove that if $\mathscr{L}$ contains a copy of $\mathscr{L}(\mathbf{R}^3)$, then $\mathscr{L}$ has no imbedding into a Boolean algebra; indeed, it follows from Stone's theorem that any Boolean algebra has $\{0,1\}$-valued probability measures, so that such imbeddings would give rise to a $\{0,1\}$-valued probability measure on $\mathscr{L}(\mathbf{R}^3)$. Z–S also proved that for arbitrary $\mathscr{L}$, if one assumes that there are imbeddings into Boolean algebras which satisfy (d) of (4) in the following stronger form

(5) (d′) $h(x \vee y) = h(x) \vee h(y)$ if $x \vee y$ exists in $\mathscr{L}$,

then $x \vee y$ will exist in $\mathscr{L}$ only when x and y commute; thus if $\mathscr{L}$ is in addition a lattice (e.g., a logic), such an imbedding will exist only when $\mathscr{L}$ is already a Boolean algebra.

The analysis of K–S is along similar lines but it is formulated in the language of partial Boolean algebras. A *partial Boolean algebra* is a set L with a family $\mathfrak{B} = \{B\}$ of subsets of L with the following properties:

(6)
(a) each $B \in \mathfrak{B}$ has the structure of a Boolean algebra; if $B, B' \in \mathfrak{B}$ and $B \subset B'$ (as sets), B is a Boolean subalgebra of B';
(b) given $B, B' \in \mathfrak{B}$, $B \cap B' \in \mathfrak{B}$;
(c) each element of L is in some $B \in \mathfrak{B}$;
(d) if $a_1, \ldots, a_n \in L$ and any two of them are contained in some member of $\mathfrak{B}$, there is a member of $\mathfrak{B}$ that contains all of them.

It follows easily from these axioms that L has the following properties:

(7)
(a′) the unit and null elements of all members of $\mathfrak{B}$ are the same; we write 0 and 1 for them;
(b′) if $a \in L$ there is a unique $a^{\perp} \in L$ such that $a^{\perp}$ is the complement of a in any $B \in \mathfrak{B}$ that contains a;
(c′) write, for $a, b \in L$, $a \leqslant b$ to mean that a and b are in some $B \in \mathfrak{B}$ and satisfy this relation there; then $\leqslant$ is a partial order on L;
(d′) with respect to $(\leqslant, \perp)$, L becomes orthocomplemented and weakly modular.

Obviously, if we start with an orthocomplemented weakly modular L its Boolean subalgebras will give L the structure of a partial Boolean algebra if (d) of (6) is satisfied. A *morphism* of a partial Boolean algebra $(L, \mathfrak{B})$ into another $(L', \mathfrak{B}')$ is a map $h: L \to L'$ such that for any $B \in \mathfrak{B}$, there is a $B' \in \mathfrak{B}'$ such that $h(B) \subset B'$ and $h: B \to B'$ is a morphism of Boolean algebras. *Imbeddings* are injective morphisms.

For K–S the basic assumption is that the set $\mathscr{L}$ of experimental propositions is a partial Boolean algebra, and as we mentioned earlier, their interpretation of the existence of hidden variables will imply the existence of an imbedding of $\mathscr{L}$ into a Boolean algebra. They prove that the existence of such an imbedding is equivalent requiring that if $a, b \in \mathscr{L}$ and $a \neq b$, there is a $\{0,1\}$-valued probability measure h on $\mathscr{L}$ such that $h(a) \neq h(b)$. They then proceed to construct explicitly a finite subset F of $\mathscr{L}(\mathbf{R}^3)$ with the property that the partial Boolean subalgebra of $\mathscr{L}(\mathbf{R}^3)$ generated by F does not admit a single $\{0,1\}$-valued probability measure.

The case of the logics of the two-dimensional spaces has been completely excluded in the above discussion. The spin observables of an electron are an example of such a system. Here the Hilbert space $\mathscr{H} = \mathbf{C}^2$. If

$$(8) \qquad \sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \qquad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

are the Pauli spin matrices, then for any $\vec{a} = (a_1, a_2, a_3)$ in the unit sphere S^2 of $\mathbf{R}^3$, the self-adjoint operator $\vec{a}\cdot\vec{\sigma} = a_1\sigma_1 + a_2\sigma_2 + a_3\sigma_3$ corresponds to the spin component along the $\vec{a}$-direction; its spectral projections are $\frac{1}{2}(1 \pm \vec{a}\cdot\vec{\sigma})$, corresponding to the eigenvalues ± 1. The map $\vec{a} \rightarrow \frac{1}{2}(1 + \vec{a}\cdot\vec{\sigma})$ is a bijection of S^2 with the set of (Hermitian) one-dimensional projections of $\mathscr{H}$, and gives rise to a bijection $p \rightarrow f_p$ between the set of probability measures p on $\mathscr{L}(\mathscr{H})$ and functions $f\colon S^2 \rightarrow [0,1]$ such that

$$(9) \qquad f(\vec{a}) \geqslant 0, \qquad f(\vec{a}) + f(-\vec{a}) = 1 \qquad (\vec{a} \in S^2),$$

with $p(\frac{1}{2}(1 + \vec{a}\cdot\vec{\sigma})) = f_p(\vec{a})$. The quantum states, i.e., the states in the conventional model of Quantum Mechanics are the functions $q_{\vec{b}}$ $(\vec{b} \in S^2)$ defined by

$$(10) \qquad q_{\vec{b}}(\vec{a}) = \tfrac{1}{2}(1 + \vec{a}\cdot\vec{b}) \qquad (\cdot = \text{scalar product in } \mathbf{R}^3).$$

It is clear that the functions (10) form only a small part of the convex set of functions (9) (even if we assume they are Borel). The extreme points of the latter set are the functions that take only the values 0 and 1, the dispersion free states; they are the characteristic functions 1_A where $A \subset S^2$ is any subset with $A \cup (-A) = S^2$, $A \cap (-A) = \varnothing$. This suggests that we can express the states (10) as mixtures of suitable families of dispersion free states and hence that a classical imbedding may exist for the quantum system associated to $\mathscr{H}$. This is in fact so, and the details were first given in articles of Kochen-Specker (loc. cit.) and J. S. Bell (*On the problem of hidden variables in Quantum Mechanics*, Rev. Mod. Phys., **38** (1966), pp. 447–452). In what follows we shall describe briefly the construction of Kochen-Specker; Bell's example is built along similar lines.

The space of "hidden variables" is S^2. Its points $\vec{w}$ define dispersion free states $f_{\vec{w}}$; for each self-adjoint operator $A \in B(\mathscr{H})$, there corresponds a function F_A on S^2; and for each pure quantum state q, there corresponds a probability measure μ_q on S^2. The F_A and μ_q are related by

$$(11) \qquad \begin{aligned} &\text{(i)} \quad F_{u(A)} = u(F_A) \quad \text{for any Borel function } u, \\ &\text{(ii)} \quad q(A) = \int_{S^2} F_A d\mu_q. \end{aligned}$$

The relation (i) asserts that the functional relationships among quantum observables are preserved under the map $A \rightarrow F_A$, and relation (ii) asserts that the probability distribution of A in the state q is that of F_A with respect to μ_q. If $A = \frac{1}{2}(\lambda I + \vec{r}\cdot\vec{\sigma})$ $(\lambda \in \mathbf{R})$ (having eigenvalues $\frac{1}{2}(\lambda \pm 1)$),

$$F_A = \tfrac{1}{2}(\lambda+1)\, 1_{\vec{w}\cdot\vec{r}\geqslant 0} + \tfrac{1}{2}(\lambda-1)\, 1_{\vec{w}\cdot\vec{r}<0}, \tag{12}$$

where $1_{\vec{w}\cdot\vec{r}\geqslant 0}$ is the characteristic function of the set of $\vec{w}$'s with $\vec{w}\cdot\vec{r}\geqslant 0$, and so on. It is immediate that the assignment (12) possesses the property (i) of (11). If $q_{\vec{b}}$ is the quantum state (10), the associated probability measure $\mu_{\vec{b}} = \mu_{q_{\vec{b}}}$ is defined by

$$d\mu_{\vec{b}} = 4(\vec{b}\cdot\vec{w})\, 1_{\vec{b}\cdot\vec{w}\geqslant 0}\, d\vec{w}, \tag{13}$$

where $d\bar{w}$ is the surface measure on S^2, given by $d\vec{w} = (1/4\pi)\sin\theta d\theta\, d\varphi$ in the polar coordinates

$$\vec{w} = (\sin\theta\cos\varphi, \sin\theta\sin\varphi, \cos\varphi)\, (0 \leqslant \theta < \pi, 0 \leqslant \varphi < 2\pi).$$

The relation (ii) of (11) then follows from the formula

$$\frac{1+\vec{b}\cdot\vec{r}}{2} = \int_{\vec{w}\cdot\vec{r}>0,\, \vec{w}\cdot\vec{b}>0} 4(\vec{b}\cdot\vec{w})\, d\vec{w}. \tag{14}$$

Finally, for each $\vec{w}$ let $f_{\vec{w}}(\vec{a} \to f_{\vec{w}}(\vec{a}))$ be the function which is 1 when $\vec{w}\cdot\vec{a} > 0$, 0 when $\vec{w}\cdot\vec{a} < 0$, and which is defined so satisfy (9) when $\vec{w}\cdot\vec{a} = 0$. Then $f_{\vec{w}}$ defines a dispersion free state on $\mathscr{L}(\mathscr{H})$, and the formula (14) becomes

$$q_{\vec{b}} = \int_{S^2} f_{\vec{w}}\, d\mu_{\vec{b}}(\vec{w}) \tag{15}$$

exhibiting $q_{\bar{b}}$ as a mixture of dispersion free states; the absence of explicit definition of $f_{\vec{w}}(\vec{r})$ when $\vec{w}\cdot\vec{r} = 0$ does not matter, since this set has measure 0 in $\vec{w}$.

We now turn our attention to the analysis from the hidden variables point of view, due to Bell, of the Einstein–Podolsky–Rosen example (J.S. Bell, *On the Einstein–Podolsky–Rosen paradox*, Physics, **1** (1964), pp. 195–200); actually Bell considers the variant, due to Bohm and Aharanov, of the EPR example. The system consists of a pair of spin $\frac{1}{2}$ particles and only their spin observables are of interest. Thus the Hilbert space is (cf. the remarks below in the discussion of composite systems)

$$\mathscr{H} = \mathbf{C}^2 \otimes \mathbf{C}^2.$$

The particles are in the "singlet" state defined by the state vector

$$\Phi = \frac{1}{\sqrt{2}}(\varphi^+ \otimes \varphi^- - \varphi^- \otimes \varphi^+), \tag{16}$$

where $\varphi^{\pm}$ are the spin states of eigenvalues ± 1, say, in the z-direction. For instance, the spin parts of the states of the two electrons in the Helium atom or the H_2-molecule are in this state (which is the unique rotation invariant antisymmetric state); even if the electrons are stripped away by dipole radiation and become widely separated spatially, (16) describes to a good approximation the compound state of the two electrons (cf. the discussion of E. P. Wigner in *Interpretation of Quantum Mechanics*, in *Quantum Theory and Measurement*, edited by J. A. Wheeler and W. H. Zurek, Princeton University Press, Princeton, NJ, 1983, pp. 260–314, especially pp. 291–293). Let $X_{1.\vec{a}}$ and $X_{2.\vec{a}}$ be the observables measuring the spin components of the first and second particle in the direction $\vec{a}$. In Quantum Mechanics they are represented respectively by the operators $\vec{a}\cdot\vec{\sigma}\otimes 1$ and $1\otimes\vec{a}\cdot\vec{\sigma}$. Let us write $P_{\mathrm{QM}}\,(\vec{a},\vec{b})$ for the expectation value of $X_{1.\vec{a}}\cdot X_{2.\vec{b}}$ under the state Φ:

$$P_{\mathrm{QM}}(\vec{a},\vec{b}) = ((\vec{a}\cdot\vec{\sigma}\otimes\vec{b}\cdot\vec{\sigma})\,\Phi,\Phi).$$

A simple calculation gives

$$P_{\mathrm{QM}}(\vec{a},\vec{b}) = -\vec{a}\cdot\vec{b} = -\cos\theta, \tag{17}$$

where θ is the angle between $\vec{a}$ and $\vec{b}$. The question studied in loc. cit. by Bell is whether the correlated statistics of the two particles in the singlet state that is summarized by (17) can be obtained from a hidden variables model that is local in the sense introduced by Einstein, Podolsky, and Rosen in their famous paper, namely, measurements on either particle do not disturb the other. Bell's analysis showed that the correlations arising from any local hidden variable model must satisfy certain limitations that conflict with (17). For nonlocal models, however, these inequalities may be violated (cf. M. Flato *et al.*, Helv. Phys. Acta, **48** (1975), pp. 219–225).

The model consists of a probability space X of points λ, and a probability measure $d\lambda$ on X; moreover, the assumption of locality enables us to say that there are random variables $A_1(\vec{a}:\lambda)$ and $A_2(\vec{b}:\lambda)$ corresponding to $X_{1.\vec{a}}$ and $X_{2.\vec{b}}$, respectively, for $\vec{a}$, $\vec{b}\in S^2$. The $A_j(\vec{a}:\lambda)$ are ± 1 valued. The correlations are defined by

$$P_{\mathrm{HV}}(\vec{a},\vec{b}) = \int_X A_1(\vec{a}:\lambda)\,A_2(\vec{b}:\lambda)\,d\lambda,$$

where the suffix stands for hidden variables. If $\vec{a}$, $\vec{b}$, $\vec{a}'$, $\vec{b}'$ are four points of S^2 then (suppressing λ)

$$\begin{aligned} |P_{\mathrm{HV}}\,(\vec{a},\vec{b}) - P_{\mathrm{HV}}\,(\vec{a},\vec{b}')| &\leqslant \left|\int A_1(\vec{a})\,A_2(\vec{b})\,(1\pm A_1(\vec{a}')\,A_2(\vec{b}'))\,d\lambda\right| \\ &\quad + \left|\int A_1(\vec{a})\,A_2(\vec{b}')\,(1\mp A_1(\vec{a}')\,A_2(\vec{b}))\,d\lambda\right| \\ &\leqslant (1\pm P_{\mathrm{HV}}\,(\vec{a}',\vec{b}')) + (1\mp P_{\mathrm{HV}}\,(\vec{a}',\vec{b}). \end{aligned}$$

So (cf. M. Lamehi-Rachti and W. Mittig, Phys. Rev. D, **14** (1976), pp. 2543–2555) we obtain the inequality

$$(18) \qquad |P_{\mathrm{HV}}(\vec{a},\vec{b}) - P_{\mathrm{HV}}(\vec{a},\vec{b}')| + |P_{\mathrm{HV}}(\vec{a}',\vec{b}') + P_{\mathrm{HV}}(\vec{a}',\vec{b})| \leqslant 2,$$

which must be satisfied by the correlations in any local hidden variables model. In Quantum Mechanics, $P_{\mathrm{QM}}(\vec{a},\vec{a}) = -1$; P_{HV} satisfies this if and only if

$$(19) \qquad A_1(\vec{a}:\lambda) = -A_2(\vec{a}:\lambda).$$

If we assume (19), then (18) becomes

$$(20) \qquad |P_{\mathrm{HV}}(\vec{a},\vec{b}) - P_{\mathrm{HV}}(\vec{a},\vec{c})| \leqslant 1 + P_{\mathrm{HV}}(\vec{b},\vec{c}).$$

This is *Bell's inequality*. It is trivial to check that (17) violates (20). Indeed, if θ_{12}, θ_{23}, θ_{31} are the angles between $\vec{a}$, $\vec{b}$, $\vec{b}$, $\vec{c}$, and $\vec{c}$, $\vec{a}$, then for (17) to obey (20) we should have

$$\sin^2(\theta_{12}/2) + \sin^2(\theta_{23}/2) \geqslant \sin^2(\theta_{31}/2),$$

which is false if $\theta_{12} = \theta_{23} = \pi/3$, $\theta_{31} = 2\pi/3$.

If we assume the model to have natural covariance properties with respect to rotations, then $P_{\mathrm{HV}}(\vec{a},\vec{b})$ depends only on the angle θ between $\vec{a}$ and $\vec{b}$ say $P_{\mathrm{HV}}(\vec{a},\vec{b}) = P_{\mathrm{HV}}(\theta)$. Obviously,

$$P_{\mathrm{HV}}(\theta) = P_{\mathrm{HV}}(-\theta) = -P_{\mathrm{HV}}(\theta+\pi).$$

The inequalities (18) now imply bounds for $P_{\mathrm{HV}}(\theta)$. Indeed, taking $\vec{a}'$, $\vec{b}'$ coplanar with $\vec{a}$, $\vec{b}$ we get, writing P for P_{HV},

$$(21) \qquad |P(\theta) - P(\theta+\alpha)| + |P(\theta') + P(\theta'+\alpha)| \leqslant 2,$$

for all θ, θ', α from which one gets (cf. Lamehi-Rachi and Mittig, loc. cit.),

$$(22) \qquad |P(\pi/6)| \leqslant \tfrac{2}{3}, \quad |P(\pi/4)| \leqslant \tfrac{1}{2}, \quad |3P(\pi/3) + P(0)| \leqslant 2.$$

The article of Bell attracted the attention of physicists and led to experiments whether (22) or (17) is true (cf. Wigner loc. cit.). The experimental results so far support Quantum Mechanics (see some of the experimental evidence presented in the articles of the collection edited by Wheeler and Zurek, loc. cit.).

Even though the basic Hilbert space has dimension 4 here, the above treatment of the hidden variables question cannot be subsumed under the von Neumann–Mackey–Gleason paradigm. The point is that for each $\lambda \in X$, the model gives only the values of $\mathrm{Exp}(X_{1,\vec{a}}{}^m \cdot X_{2,\vec{b}}{}^n)$ $(m, n \geqslant 0)$; and hence each $\lambda \in X$ can define a dispersion free state only on the maximal abelian subalgebras $\mathscr{A}(\vec{a},\vec{b})$ of $\mathscr{H}$ generated by $\vec{a}\cdot\vec{\sigma}\otimes 1$ and $1\otimes\vec{b}\cdot\vec{\sigma}$. Hidden variable theories such as these are called *contextual* in the book of Beltrametti and Cassinelli (cf. notes on Chapter III). The reader who wants to know more about these and get a sharper perspective on the entire question of hidden variables cannot do better than consult the book of Beltrametti and Cassinelli, for its thorough treatment of related issues as well as the references it contains to other literature.

For a question as deep as the one that asks whether quantum mechanical description of the atomic world is complete, a mere mathematical analysis will hardly provide a completely satisfying answer. The argument must also be based on analyzing possible measurements and the limitations imposed on them by the uncertainty relations. The paper of Einstein, Podolsky, and Rosen, and Bohr's refutation of it, offer perhaps the most profound discussion of this question from the natural philosopher's point of view (cf. the relevant articles in the collection edited by Wheeler and Zurek, loc. cit.).

2. *Algebraic formulation of Quantum Mechanics.* In some sense the algebraic method of viewing the foundations of Quantum Mechanics is the oldest of all approaches, and arguably, the most flexible one. The early papers of Born, Jordan, Heisenberg, and Dirac repeatedly emphasized that quantum observables are represented by matrices and so form a noncommutative algebra, in sharp contrast to the commutative algebra of classical observables (for a historical account as well as translations of the early papers, see *Sources of Quantum Mechanics*, B. L. van der Waerden, Dover, New York, 1967). Eventually this came to be formalized as the principle that the quantum observables are in one–one correspondence with the self-adjoint operators on a (separable, complex) Hilbert space. In spite of the overwhelming success of the models built out of this assumption it was always understood that this was a very *ad hoc* solution to the mathematical, physical, and philosophical problems raised by atomic physics. Especially troublesome was the fact that both the operations of addition and multiplication involved possibly noncommuting (hence nonsimultaneously observable) observables, thus making it very difficult to give any convincing "explanation" for them. (Perhaps one should not put too much faith in such "explanations". Most explanations make implicit or explicit use of classical perceptions and are not too relevant in the atomic realm.) There was also the fact that every self-adjoint operator was supposed to represent an observable; in the presence of superposition rules one should modify this suitably, but even then this was gradually seen as a far-reaching enlargement of what can be effectively observed in atomic systems—see *Interpretation of Quantum Mechanics*, E. P. Wigner, in *Quantum Theory and Measurement*, edited by J. A. Wheeler and W. H. Zurek, p. 298. Multiplication was felt to present the greater problem since it could be argued that if A and B are observables, $A+B$ can be effectively defined by prescribing that its expected value in any state is the sum of the expected values of A and B in that state. This, of course, is a little deceptive, for, as we have

already remarked above, the linearity of expectation values is a deep property of the propositional logic; its validity for the standard logics (in dimension $\geqslant 3$) is a consequence of Gleason's theorem while its truth for the more general logics of projections in von Neumann algebras has only recently been proved (see the remarks on Gleason's theorem, and also §14.5 of the book of Beltrametti and Cassinelli, cited in the notes to Chapter III).

The first systematic investigations on the algebraic structure of the set of observables are due to Jordan, von Neumann, and Wigner; the initial papers are those of Jordan in Gött. Nachr. (1932), pp. 569–575, Gött. Nachr. (1933), pp. 209–217, and Z. Phys., **80** (1933), pp. 285–291, and that of Jordan, von Neumann, and Wigner in Ann. Math., **35** (1934), pp. 29–64. One of the motivations for these studies was the feeling that the generalizations of Quantum Mechanics to relativistic and nuclear physics might force such a more general view of observables. Jordan's basic observation was that although the set of bounded self-adjoint operators do not form an algebra under multiplication, it is nevertheless an algebra over **R**, even abelian, if we define a new product by $A\cdot B=\frac{1}{2}(AB+BA)$. This product is bilinear and satisfies the distributive laws $A\cdot(B+C)=A\cdot B+A\cdot C$, although not the associative law; moreover, the identity $A\cdot B=\frac{1}{2}[(A+B)^2-A^2-B^2]$ shows that this product can be obtained from the linear and power structures. Nowadays such algebras, emerging as they did out of the pioneering studies of Jordan, are known as Jordan algebras; see the book of Jacobson, *The structure of representations of Jordan algebras*, Amer. Math. Soc. Colloq. Publications, vol. XXXIX, 1968, Providence, R.I.

Typically, one starts with an associative but not necessarily commutative algebra and define the new multiplication by $a\cdot b=\frac{1}{2}(ab+ba)$ to obtain a Jordan algebra. But it is not always possible to obtain a given Jordan algebra in this fashion. The fundamental result of Jordan, von Neumann, and Wigner in their 1934 paper is the complete classification of all finite dimensional Jordan algebras A over **R** which are irreducible and *formally real*, i.e., satisfy

$$a_1,\dots,a_n\in A,\qquad a_1{}^2+\dots+a_n{}^2=0\ \Rightarrow\ a_1=\dots=a_n=0.$$

It follows from their classification that, with one exception (that of $M_3{}^8$, the algebra of self-adjoint 3×3 matrices over the Cayley numbers), all of these are obtained in the above-mentioned manner from associative algebras; and further, that if some special cases are disregarded, one gets just the Jordan algebras of Hermitian matrices over **R**, **C**, or **H**.

The finite dimensionality assumption is a very severe one and von Neumann began the study of infinite dimensional topological algebras (Mat. Sb., **1** (1936), pp. 415–484; *Collected Works*, Vol. III, No. 9, Pergamon, Oxford, 1961). His investigation in this paper contains a deep analysis of the spectral theory of individual elements and its relationship to simultaneous observability. Although he had, in connection with his analysis of the possible existence of hidden variables (cf. his book), introduced the notion of the state of a system as a positive definite linear form on the observable algebra, this paper does not contain a study of the states.

With the appearance, in the early 1940s, of the path-breaking papers of Gel'fand and Naĭmark on Banach algebras, the essential aspects of this approach became clearer. This theme was taken up by I. E. Segal in 1947 (*Postulates for general Quantum Mechanics*, Ann. Math., **48** (1947), pp. 930–948). Segal's postulates on the observable algebra were simpler and more direct than von Neumann's, and were motivated by what is operationally possible; in addition, he studied in detail the structure of the states.

The case when the (bounded) observables are the self-adjoint elements of a complex **C***-algebra (a complex Banach algebra with involution * such that $\|x^*x\| = \|x\|^2 = \|x^*\|^2$, $\|1\| = 1$) is a possible system in the sense of Segal. Although special, it is an important one since it includes the conventional model of Quantum Mechanics. If $\mathfrak{U}$ is the **C***-algebra whose self-adjoint elements are in correspondence with the bounded observables of the system, a state of the system is a linear function f on $\mathfrak{U}$ such that $f(1) = 1$ and $f(x^*x) \geqslant 0$ for all $x \in \mathfrak{U}$. The Gel'fand–Naĭmark–Segal representation theorem associates to f an essentially unique triple $(\mathscr{H}_f, \pi_f, \psi_f)$ consisting of a Hilbert space $\mathscr{H}_f$, a unit vector ψ_f in $\mathscr{H}_f$, and a *-representation π_f of $\mathfrak{U}$ in $\mathscr{H}_f$ with ψ_f as cyclic vector such that $f(x) = (\pi_f(x)\psi_f, \psi_f)$ $(x \in \mathfrak{U})$; f is pure if and only if π_f is irreducible. In this way state vectors are once again unit vectors in Hilbert spaces; but the Hilbert space is not fixed as in conventional Quantum Mechanics, but varies with the state. The importance and usefulness of these more general types of observable algebras become clear in the theory of Second Quantization where one works with systems of identical particles (photons, electrons, He^3, He^4, etc.), but where the number of particles may not remain fixed. In this case it turns out in contrast to the case of systems of a fixed number of particles, that the canonical commutation rules (or anti-commutation rules) have many essentially different representations. Nevertheless, it was proved by Segal (Mat. Fys. Medd. Dan. Vid. Selsk, **31** (1959), No. 2) that the **C***-algebra, which is the norm-closure of the observables depending only on finitely many particles, is canonically determined; therefore it may be regarded as the algebra of field observables. The possibility that one may have to vary the Hilbert space with the state has also been discussed by Dirac (*Lectures on Quantum Field Theory*, Belfer Graduate School of Science Monograph Series, No. 3, Belfer Graduate School of Science, Yeshiwa University, New York, 1966).

The study of infinite dimensional observable algebras has been pursued in recent years in another direction, closer to the original theme of Jordan, von Neumann, and Wigner. This is the study of what are called JB, JBW algebras, which are the Jordan algebra analogues of the self-adjoint parts of C*-algebras and von Neumann algebras. The state spaces of these are compact convex sets with an interesting geometric structure, and they can be characterized in an intrinsic manner. This thus opens up another approach to the foundations in Quantum Mechanics. For a survey of many aspects of this theme we refer the reader to the article of E. M. Alfsen (*On the state spaces of Jordan and C*-algebras*, in *Algebres d'operateurs et leurs applications en Physique Mathematique*, CNRS Coll. Intern. No. 274, 1979, pp. 15–40).

In recent years people (Haag, Kastler, and others) have used algebras of operators for formulating some general properties of physical systems (for example, a quantized field) extended in space-time, that are local in the sense that the observables that are localized in noncausally connected regions commute with each other. These ideas have clarified somewhat the nature of the high-energy scattering processes that dominate elementary particle physics. The reader who wants to understand the algebraic approach may start with the book of G. G. Emch (*Algebraic Methods in Statistical Mechanics and Quantum Field Theory*, Wiley, New York, 1972), and the references cited there. It should be remarked, however, that the mathematical questions arising in the attempt to construct rigorous models of scattering processes involving particle creation and annihilation (Quantum Electrodynamics, for example) are quite deep and are tied up with themes of an entirely different sort (cf. J. Glimm and A. Jaffe, *Quantum Physics: A Functional Integral Point of View*, Springer-Verlag, New York, 1981).

The approach to Quantum Mechanics via functional integrals was pioneered by R. P. Feynman [2] (cf. also his book with A. R. Hibbs, *Quantum Mechanics and Path Integrals*, McGraw-Hill, New York, 1965) and is known as the *path integral formalism*. It is widely used by physicists although mathematically difficult to justify. The attempts to understand this formalism at a rigorous level have proved very beneficial and effective in the mathematical theory of elementary particles and their interactions; see the book of Glimm and Jaffe, loc. cit. Finally, for a view of the foundations of Quantum Theory that is very close to phenomenology one should refer to the series of papers of J. Schwinger which are a part of his book *Quantum Kinematics and Dynamics*, W. A. Benjamin, New York, 1970.

3. *Statistics of mixed states.* Consider a system governed by the standard logic $\mathscr{L}(\mathscr{H})$ where $\mathscr{H}$ is a separable complex Hilbert space, and let it be in the state p_U where U is a von Neumann operator on $\mathscr{H}$. By the spectral theorem we have $U = \sum_n \lambda_n P_{L_n}$ where $\lambda_1, \lambda_2, \ldots$ are the *distinct* eigenvalues >0 of U, L_n the eigensubspace corresponding to λ_n, and P_{L_n}

is the orthogonal projection $\mathscr{H} \to L_n$. If $d_n = \dim(L_n)$, $w_n = d_n \lambda_n$, and $U_n = d_n^{-1} P_{L_n}$, then $\operatorname{tr}(U_n) = 1$, $\sum_n w_n = 1$, and the formula

$$p_U = \sum_n w_n p_{U_n} \tag{23}$$

exhibits p_U as a canonical mixture of the states p_{U_n} with weights w_n, the supports of the p_{U_n} being mutually orthogonal. If U has a simple spectrum, i.e., if $d_n = 1$ for all n, the p_{U_n} are all pure states; and (23) is the *unique* decomposition of p_U as a convex combination of pure states defined by orthonormal vectors. If some of the d_n are > 1, the right side of (23) can be split further but not uniquely; indeed, if $(u_{ni})_{1 \leqslant i \leqslant d_n}$ is any orthonormal basis of L_n we have

$$p_U = \sum_n w_n \sum_{1 \leqslant i \leqslant d_n} d_n^{-1} p_{u_{ni}}, \tag{24}$$

which exhibits p_U as a convex combination of pure states defined by the orthonormal family of vectors (u_{ni}). Thus, as soon as some $d_n > 1$ we cannot definitely say of which pure states p_U is a mixture of; the phenomenon is clearly nonclassical and is another example of the sharp contrast between quantum and classical statistics.

4. *Composite systems.* One of the most widely used methods of analyzing complicated systems is that of viewing them in terms of suitable subsystems and their interactions. For example, for many purposes it is completely adequate to view an atom as a system of electrons moving in a centrally symmetric force field. We are thus led to the general problem of composition of systems. Let $\mathfrak{S}_1, \mathfrak{S}_2, \ldots, \mathfrak{S}_N$ be quantum mechanical systems; in conventional Quantum Mechanics we associate complex separable Hilbert spaces $\mathscr{H}_1, \ldots, \mathscr{H}_N$, to them with the usual prescriptions of states and observables. It is then possible to form a composite system $\mathfrak{S}$ containing $\mathfrak{S}_j$ as subsystems in a natural manner, and in fact to do this in a universal way. The logic of $\mathfrak{S}$ is $\mathscr{L}(\mathscr{H})$ where

$$\mathscr{H} = \mathscr{H}_1 \otimes \cdots \otimes \mathscr{H}_N.$$

The map

$$i_\alpha : A \to 1 \otimes \cdots \otimes A \otimes \cdots \otimes 1 \quad (1 \leqslant \alpha \leqslant N, A \text{ in the } \alpha\text{th factor})$$

imbeds the algebra $B(\mathscr{H}_\alpha)$ in $B(\mathscr{H})$; by restricting the A's to the projection operators, each i_α gives rise to an imbedding

$$i_\alpha : \mathscr{L}(\mathscr{H}_\alpha) \hookrightarrow \mathscr{L}(\mathscr{H}) \qquad (1 \leqslant \alpha \leqslant N).$$

Thus the properties of the $\mathfrak{S}_\alpha$ are faithfully reflected within $\mathfrak{S}$. If σ is a state of $\mathfrak{S}$, i.e., a probability measure on $\mathscr{L}(\mathscr{H})$, then

$$\sigma_\alpha = \sigma \circ i_\alpha \tag{25}$$

is a probability measure on $\mathscr{L}(\mathscr{H}_\alpha)$ and so defines a state of $\mathfrak{S}_\alpha$, *the state that* $\mathfrak{S}_\alpha$ *is in when* $\mathfrak{S}$ *is in the state* σ. The map

$$(26) \qquad \sigma \to (\sigma_1, \ldots, \sigma_N)$$

is clearly surjective; if $\sigma_\alpha{}' = p_{U_\alpha}$ where U_α is a von Neumann operator on $\mathscr{H}_\alpha$ and $\tau = p_U$ where $U = U_1 \otimes \ldots \otimes U_N$, then $\tau \circ i_\alpha = \sigma_\alpha{}'$ for all α. In general the σ_α do not determine σ uniquely; it can be proved that this is so if and only if at most one of the σ_α is nonpure (cf. von Neumann [1], pp. 425–429); in particular, if $u_\alpha \in \mathscr{H}_\alpha$ is a unit vector, $\sigma_\alpha = p_{u_\alpha}$, and $u = u_1 \otimes \ldots \otimes u_N$, p_u is the unique state of $\mathscr{L}(\mathscr{H})$ such that $p_u \circ i_\alpha = p_{u_\alpha}$ for all α. However, if $u \in \mathscr{H}$ is a unit vector which is not in this "split" form, some of the states $p_u \circ i_\alpha$ will not be pure, thereby illustrating the difference between the quantum and classical compositions.

As a simple example let $N = 2$, $\mathscr{H}_1 = \mathscr{H}_2 = \mathbf{C}^2$; we can think of each $\mathfrak{S}_j$ as a pure spin system of spin $\frac{1}{2}$. If $\{\varphi^+, \varphi^-\}$ is an orthonormal basis in $\mathbf{C}^2$ which corresponds to the ± 1 values of the spin in some (fixed) direction in space,

$$(27) \qquad \Phi = \frac{1}{\sqrt{2}}(\varphi^+ \otimes \varphi^- - \varphi^- \otimes \varphi^+)$$

represents the so-called "spin singlet" state of the particle pair. The states $\sigma_\alpha = p_\Phi \circ i_\alpha$ $(\alpha = 1, 2)$ are then both mixed, being equal to $\frac{1}{2}(p_{\varphi^+} + p_{\varphi^-})$. The probability of having spin parallel to any given direction is $\frac{1}{2}$ for each of the particles; but the spins are so correlated that, given that the first particle has spin along the direction $\vec{a}$, the probability is 1 that the second particle has spin along $-\vec{a}$. This example is a variant due to Bohm and Aharanov (Phys. Rev., **108** (1957), p. 1070) of the example used by Einstein, Podolsky, and Rosen in their famous paper (Phys. Rev., **47** (1935), pp. 777–780). By measuring the spin of the first particle in any direction $\vec{a}$ we can predict with certainty the spin of the particle along $\vec{a}$ without apparently disturbing its state, a paradoxical result since the spin components along different directions are not simultaneously observable. The paradox arises because of the assumption (which ultimately involves preconceptions about the structure of matter, among other things) that a measurement of the spin state of the first particle does not affect the state of the second when they are spatially far apart; the paradox disappears if one realizes that when the composite system is in a state such as (27), a measurement of the spin state of the first particle changes the state of the composite system, and consequently also the state of the second system (cf. the detailed analysis of the situation in Bohr's reply to the EPR paper, Phys. Rev., **48** (1935), pp. 696–702).

The theory of composition of systems in contexts other than the standard ones is more involved. For instance, since there is no reasonable way to form tensor products of vector spaces over noncommutative division rings, the theory of composite systems in quaternion Quantum Mechanics becomes more difficult. For quaternionic Quantum Mechanics, the impossibility of forming tensor products is interpreted by Finkelstein *et al.* [1] to mean that *there are no truly independent systems*, i.e., that there are complementarity relations between observables of any two systems in quaternionic Quantum Mechanics. The issue becomes even more complicated when we treat it in the framework of general logics. The reader is referred to the discussion in Chapter 24 of the book of Beltrametti and Cassinelli (cf. Notes on Chapter III) and to the literature cited there.

The above considerations have to be modified substantially when one wants to treat systems of *identical* particles, because, in Quantum Mechanics there are no experiments that can keep track of individual particles throughout the course of some physical process. Thus, in such assemblies of identical particles, only the physical quantities that involve the various particles *symmetrically* are permitted to be observables. Mathematically, this means the following: if $\mathscr{H}$ is the Hilbert space of a single particle, then for a system of N such particles, the self-adjoint operators on

$$\mathscr{H}^{(N)} = \mathscr{H} \otimes \dots \otimes \mathscr{H} \quad (N \text{ times})$$

that represent physical observables must commute with the natural action of the permutation group S_N of $\{1, \dots, N\}$ on $\mathscr{H}^{(N)}$; here we recall that any $\sigma \in S_N$ acts on $\mathscr{H}^{(N)}$ by sending $u_1 \otimes \dots \otimes u_N$ to

$$u_{\sigma^{-1}(1)} \otimes \dots \otimes u_{\sigma^{-1}(N)} \, (u_i \in \mathscr{H}).$$

Let $\mathfrak{U}$ be the algebra of all operators on $\mathscr{H}^{(N)}$ commuting with this action on S_N. To analyze the structure of $\mathfrak{U}$ we proceed as follows. For any irreducible character τ of S_N let $E(\tau)$ be the projection operator

$$(N! \sqrt{\dim(\tau)})^{-1} \sum_{s \in S_N} \tau(s) s \quad \text{on } \mathscr{H}^{(N)}$$

that projects onto the space of "τ-symmetric" tensors; if

$$\mathscr{H}^{(N)}(\tau) = E(\tau)\,(\mathscr{H}^{(N)}),$$

the $\mathscr{H}^{(N)}(\tau)$ are mutually orthogonal, stable under S_N and $\mathfrak{U}$ with

$$(28) \qquad \mathscr{H}^{(N)} = \bigoplus_{\tau} \mathscr{H}^{(N)}(\tau);$$

moreover, we can factorize each $\mathscr{H}^{(N)}(\tau)$ (but noncanonically), as

$$(29) \qquad \mathscr{H}^{(N)}(\tau) \approx K(\tau) \otimes L(\tau),$$

where $\mathfrak{U}$ (resp. S_N) acts only on the first (resp. second) factor in such a way that S_N acts irreducibly on $L(\tau)$ with character τ while $\mathfrak{U}$ is isomorphic to $B(K(\tau)) \otimes 1$ (see Weyl [1]). Thus, if $\mathscr{L}$ is the logic of S_N-stable closed subspaces of $\mathscr{H}^{(N)}$, then the $\mathscr{H}^{(N)}(\tau)$ are the atoms of the center of $\mathscr{L}$, the logics $\mathscr{L}(\tau) = \mathscr{L} \cap \mathscr{H}^{N)}(\tau)$ are standard (being isomorphic to $\mathscr{L}(K(\tau))$, and

$\mathscr{L} = \bigoplus_\tau \mathscr{L}(\tau)$ is the central decomposition of $\mathscr{L}$. Although (29) is not canonical, once we fix such a factorization, for each τ, the pure states of $\mathscr{L}$ may be identified with the rays of $\mathscr{H}^{(N)}$ that lie in some $K(\tau)$, so that the $K(\tau)$ are the *coherent subspaces* of pure states. The reader should view this as an example of noncommuting superselection rules.

Two specific coherent spaces of pure states deserve special attention. These are the cases where τ is one-dimensional so that $K(\tau) = \mathscr{H}^{(N)}(\tau)$. More precisely, we consider the two cases: (a) $\tau = 1$, $\tau(s) = 1$ for all $s \in S_N$; (b) $\tau = \text{sign}$, $\tau(s) = \text{signature of } s$ for all $s \in S_N$. In (a), $\mathscr{H}^{(N)}(\tau)$ is the subspace of symmetric tensors of rank N over $\mathscr{H}$, or the subspace of degree N, $S^{(N)}(\mathscr{H})$, in the *symmetric algebra* $S(\mathscr{H})$ over $\mathscr{H}$. In (b), $\mathscr{H}^{(N)}(\tau)$ is the subspace of skew-symmetric or *alternating* tensors of rank N over $\mathscr{H}$, or the subspace of degree N, $\Lambda^{(N)}(\mathscr{H})$, in the *Grassmann or Exterior algebra* $\Lambda(\mathscr{H})$ over $\mathscr{H}$. In atomic theory, in order to get agreement with experiment it is necessary to assume that for systems of N electrons no states other than those in $\Lambda^{(N)}(\mathscr{H})$ occur in nature, i.e., the logic is the standard one associated with the Hilbert space $\Lambda^{(N)}(\mathscr{H})$. If u_i $(1 \leqslant i \leqslant N)$ are orthonormal vectors in $\mathscr{H}$, it is customary to interpret $u_1 \wedge \ldots \wedge u_N$ as "the state in which the electrons (in some order) occupy the states $u_1, \ldots, u_N$". The fact that $u_{i_1} \wedge \ldots \wedge u_{i_N} = 0$ if two of the i_μ's are identical is usually formulated as the *Pauli exclusion principle*, namely, "no two electrons can be in the same state." In radiation theory which studies photon assemblies it becomes necessary to assume that for a system of N photons only the symmetric tensor states occur in nature, i.e., the logic is the standard one associated with the Hilbert space $S^{(N)}(\mathscr{H})$. In either case it is necessary to give additional prescriptions, typically by introducing specific operators, to describe physical quantities and processes characteristic of the systems of particles in question. The systematic way of doing this is known as *Second Quantization*.

In particle physics it is assumed that the above description in terms of $\Lambda^{(N)}(\mathscr{H})$ or $S^{(N)}(\mathscr{H})$ extends to assemblies of particles of any type whatsoever. Particles which require the exterior algebra description are known as *Fermions* and are said to obey *Fermi–Dirac statistics*, while the others are known as *Bosons* and obey *Bose–Einstein statistics*. Experiment has shown that particles with half integral spin (electrons, neutrons, protons, He^3 atom, etc.) are Fermions while particles with integral spin (photons, He^4 atom) are Bosons. All theoretical explanations of this remarkable connection between spin and statistics depend on the theory of relativistic quantum fields.

5. *Measurement in Quantum Mechanics.* The theory of measurement is at the heart of Quantum Mechanics. Indeed it is the uncontrollable disturbance caused by measurements on the physical systems being observed that leads to the statistical character of the quantum mechanical picture of the

physical world. It is natural to try to determine the extent to which one can understand the nature of the measurement process *within the framework of Quantum Mechanics itself.* Of course Quantum Mechanics gives unambiguous prescriptions for calculating the probability distribution of any physical quantity at a given instant of time, if the state of the physical system is given at an earlier time (Schrödinger's time-development equation), so that, no theory of measurement is needed to ensure the logical consistency of the theory. But the state vectors are not physical objects; they are idealizations, and one has conceptual access to them only through results of experiments. Thus, if the theory is to be confronted with experience, and is to be used as a guide for understanding existing phenomena as well as for predicting new ones, it has to contain instructions that tell us how to determine the states from physical data, i.e., how measurements on a system change its state.

The fundamental work on the general theory of measurement is that of von Neumann, most of which is contained in Chapters V and VI of his book [1]. Since then, the subject has received a lot of attention and new themes have been introduced. We shall limit ourselves to a brief discussion of von Neumann's work and some of the later developments inspired by it. The reader who wants to get a historical perspective and a deeper insight into these questions should refer to the volume edited by Wheeler and Zurek (*Quantum Theory and Measurement*, Princeton University Press, Princeton, NJ, 1983) that collects together most of the crucial papers on the subject.

For simplicity we shall restrict ourselves to the conventional model of Quantum Mechanics. Let a quantum mechanical system be in a pure state represented by the normalized vector φ (of the Hilbert space $\mathscr{H}$ underlying the system), and let A be a self-adjoint operator with a purely discrete simple spectrum that represents some observable. If $a_1, a_2, \ldots$ are the eigenvalues of A and $\varphi_1, \varphi_2, \ldots$ are the corresponding normalized eigenvectors, then a measurement (instantaneous) of A will yield the values $a_1, a_2, \ldots$ with probabilities $|(\varphi,\varphi_1)|^2$, $|(\varphi,\varphi_2)|^2, \ldots$. However, if a particular value, say a_k, is obtained, and the measurement of A is again carried out *immediately after*, the value does not change. This is interpreted in the following way: the act of measurement has changed the state of the system from φ to φ_k. Since the value a_k is obtainable only with probability $|(\varphi,\varphi_k)|^2$, one may then summarize this set of circumstances by saying that the statistical effect of the measurement of A consists in changing the state of the system from the pure state φ to the *mixed state* in which the pure states $\varphi_1, \varphi_2, \ldots$ occur with probabilities $|(\varphi,\varphi_1)|^2$, $|(\varphi,\varphi_2)|^2, \ldots$. If the initial state is not pure but mixed and U is the von Neumann operator representing it, the operator of the changed state is

$$U_A = \sum_{n \geqslant 1} (U\varphi_n, \varphi_n) P_{[\varphi_n]}, \tag{30}$$

where $P_{[\varphi_n]}$ is the orthogonal projection on the one-dimensional span of φ_n (for the discussion of the Compton–Simons experiment of light scattering by electrons on which this assumption is based, see von Neumann [1], pp. 212–214). Thus the act of measuring A induces the map

$$M_A : U \to U_A \tag{31}$$

on the convex set of states. It is clear from (30) that M_A is a convex map; but it is a convex *endomorphism* and not a convex *automorphism* since in general it changes pure states into mixed ones. The essential content of von Neumann's work is a profound analysis of the mathematical and physical nature of the endomorphisms of the type M_A.

The first question is to what extent M_A fails to be invertible. Von Neumann viewed this from the more general point of view of deciding whether (the effect of) M_A is irreversible. To this end, he introduced also the *automorphisms* of the convex set of states of the form

$$a_D : U \to DUD^{-1}; \tag{32}$$

these are induced by unitary operators D of $\mathscr{H}$ that represent the dynamical transformations. He then formulated the irreversibility of M_A as the fact that it is impossible to go from $M_A(U)$ to U by repeated applications of transformations of the form M_B, a_D. For proving this he introduced, for any U, its *entropy* $H(U)$:

$$H(U) = -\kappa\, \mathrm{tr}(U \ln U), \quad \kappa > 0 \text{ being Boltzmann's constant.} \tag{33}$$

Then he proved that

$$H(U) \geqslant 0, \quad H(U) = 0 \Leftrightarrow U = p_\varphi \quad \text{for some } \varphi \in \mathscr{H} \text{ with } \|\varphi\| = 1, \tag{34}$$

and further that

$$H(a_D(U)) = H(U), \tag{35}$$

$$H(M_A(U)) > H(U) \quad \text{unless} \quad M_A(U) = U. \tag{36}$$

The irreversibility of the transition $U \to U_A$ follows from (35), (36).

As remarked by von Neumann himself, there are many unitary invariants besides $H(U)$ that do not decrease under the measurement transition $U \to U_A$. For instance, the function that assigns to U the value $-\lambda$ where λ is the largest eigenvalue of U, has this property. But as we shall explain briefly below, the entropy function H has *thermodynamic* significance; and the relations (35)–(36) show that the measurement transition $U \to U_A$ is actually irreversible *even at the thermodynamic level.* Since it is easy to construct examples of changes $U \to U'$ where the entropy stays the same but the largest eigenvalue decreases strictly, we see that there are thermodynamically *reversible* changes which cannot be brought about by repeated applications of M_B, a_D. Thus the inequalities (36) go far beyond merely establishing the mathematical irreversibility of the transitions $U \to U_A$.

Von Neumann was led to the definition (33) of entropy by his analysis of ensembles of noninteracting systems each of which is in the state represented by the (von Neumann) operator U. He used the methods of phenomenological thermodynamics for his purpose. The only difference between the classical setting and the present one is that the individual systems are now governed by the laws of Quantum Mechanics. The problem was to calculate, for an ensemble of N systems in the state U, the decrease in entropy that should be achieved to bring the systems to a pure state p_{φ}. He found that the pure states are transformable among themselves without any change of "heat energy" and so all ensembles where the individual systems are in a pure state p_q have the same entropy, which can be normalized to be 0. He then showed, assuming the validity of the first and second laws of thermodynamics, that the ensemble of N systems with state U requires a decrease in entropy equal to $-N\kappa \operatorname{tr}(U \ln U)$, κ being Boltzmann's constant, in order that it may be transformed to the pure state p_q. Thus the entropy of the original ensemble is

$$-N\kappa \operatorname{tr}(U \ln U). \tag{37}$$

The reversal of the processes used here shows that two ensembles with the same entropy can be transformed into each other by thermodynamically admissible processes. This analysis, which led to the formula (37) for the entropy, also suggested the properties (33)–(36).

The relation (35) asserts that the entropy remains constant during time evolution if no measurements are carried out. This will appear paradoxical because classically the entropy always increases. The paradox disappears if we realize that the entropy $H(U)$ is the entropy of the microstate. The classical entropy is a macroscopic quantity and its time variations are due to the fact that the observer does not know the microstate of the system. To define an analogue of the classical entropy one may introduce a Boolean σ-algebra $\mathfrak{M}$ of "macroscopic projections" that form the logic of macroscopic quantities. If we assume that $\mathfrak{M}$ is atomic and its atoms $E_1, E_2, \dots$ are finite dimensional projections where

$$s_n = \dim(E_n) \geqslant 1, \tag{38}$$

then it is reasonable to define the *macroscopic entropy* $H_{\mathfrak{M}}(U)$ of the state U by

$$H(U) = -\sum_n \operatorname{tr}(UE_n) \ln \frac{\operatorname{tr}(UE_n)}{s_n}. \tag{39}$$

Von Neumann proved that

$$H_{\mathfrak{M}}(U) \geqslant H(U), \tag{40}$$

with equality if and only if

$$U = \sum_n (\mathrm{tr}(UE_n)/s_n)E_n. \tag{41}$$

Unlike the previous case, the macroscopic entropy is always >0 for any U. Moreover, it will change under time evolution. Von Neumann investigated this as well as the associated ergodic theorems in his paper in Z. Physik, **57** (1929), pp. 30–70, to which we refer the reader for further details.

The second aspect of the work of von Neumann is the investigation of the maps M_A from the point of view of the Quantum Mechanics of the composite system formed by the system $\mathfrak{S}$ under investigation and the system $\mathfrak{S}'$ consisting of the "measuring apparatus." Let $\mathscr{H}$ and $\mathscr{H}'$ be the Hilbert spaces associated with them and let $\overline{\mathscr{H}} = \mathscr{H} \otimes \mathscr{H}'$. Let A be as before a self-adjoint operator with a purely discrete spectrum, $a_1, a_2, \ldots$, $\varphi_1, \varphi_2, \ldots$ being the corresponding eigenvectors (normalized). The act of measurement of A in $\mathfrak{S}$ may be regarded as a "temporary" insertion of an energy coupling in the composite system that allows $\mathfrak{S}'$ to interact with $\mathfrak{S}$. The time development of the composite system in the (short) time interval of measurement is then given by the dynamical group $t \to \exp(-itE)$ where E is the energy operator in $\overline{\mathscr{H}}$. If ξ is a unit vector in $\mathscr{H}'$ representing the initial state of the measuring apparatus, and φ is a unit vector in $\mathscr{H}$ representing the initial state of $\mathfrak{S}$, the state of the composite system after measurement is $\exp(-i\tau E)(\varphi \otimes \xi)$ where $\tau > 0$ is the duration of measurement. This state induces a state of $\mathfrak{S}$ which may then be regarded *as the state of* $\mathfrak{S}$ *after measuring* A. To make sure that this scheme captures at least some of the key aspects of measurement it is natural to require that the following conditions be satisfied.

(i) There is an orthonormal set $(\xi_n)_{n \geqslant 1}$ in $\mathscr{H}'$ such that the state ξ_n of $\mathfrak{S}'$ "corresponds" to the state φ_n of $\mathfrak{S}$.

(ii) For all $\varphi \in \mathscr{H}$,

$$\exp(-i\tau E)(\varphi \otimes \xi) = \sum_n c_n(\varphi)\, \varphi_n \otimes \xi_n,$$

with

$$|c_n(\varphi)| = |(\varphi, \varphi_n)|.$$

The pairing $\varphi_n \leftrightarrow \xi_n$ is the mechanism by which the observer, recording the state ξ_n of $\mathfrak{S}'$, recognizes that $\mathfrak{S}$ is in the state φ_n. The second condition shows at once that the state of $\mathfrak{S}$ induced by $\exp(-i\tau E)(\varphi \otimes \xi)$ is

$$\sum_n |(\varphi, \varphi_n)|^2 p_{\varphi_n} = M_A(p_\varphi).$$

If for a given A we can find $(\xi_n)_{n \geqslant 1}$, ξ, E, τ such that (i) and (ii) are satisfied, then it would be justifiable to view the transitions $U \to M_A(U)$ of the states of $\mathfrak{S}$, which are irreversible within $\mathfrak{S}$, as the effects on $\mathfrak{S}$ of the perfectly reversible and continuous transitions provided in the composite system by its time development. In other words, the causal and continuous change of states arising out of the dynamical evolution gives one a unified quantum mechanical way of looking at the physical world and its interactions with the observer. Von Neumann proved that this is possible; that is, for any A, and for fixed $\tau > 0$, ξ, $(\xi_n)_{n \geqslant 1}$, there is a suitable E satisfying (i) and (ii) above.

This result of von Neumann inspired much subsequent work whose focus was on the following issue: To what extent do the requirements (i) and (ii) above as well as the operator E constructed by von Neumann give a faithful picture of measurement processes encountered in the physical world? It could be argued that the description of $\mathfrak{S}'$ by the Quantum Mechanics of the logic $\mathscr{L}(\mathscr{H}')$ does not take into account the fact that very often $\mathfrak{S}'$ is "very big" compared to $\mathfrak{S}$ and so the logic of $\mathfrak{S}'$ has a rich supply of superselection rules. The requirement that measurement is of very short duration is also troublesome if these questions are to be treated relativistically. Finally, if there is additional structure in $\mathscr{H}$, $\mathscr{H}'$, and $\overline{\mathscr{H}}$, it is not clear that the E constructed would respect them. The first example of nonmeasurability of this kind was exhibited by E. Wigner (Z. Physik, **131** (1952), p. 101). He showed that if there are *additive* conservation laws for E, then A cannot be measured unless it commutes with the conserved quantity; and further, that for approximate measurements of A with arbitrary accuracy to be possible, the measuring apparatus should be large enough to admit states which are superpositions of sufficiently many states with different quantum numbers of the conserved quantity. Actually for Wigner, A was the x-composnent of the spin of a spin $\frac{1}{2}$ particle, the conserved quantity being the z-component of the angular momentum. These results were proved in full generality by H. Araki and M. Yanase (Phys. Rev., **120** (1960), pp. 622–626). They proved that if L_1 (resp. L_2) is a bounded self-adjoint operator on $\mathscr{H}$ (resp. $\mathscr{H}'$), and if $\tau > 0$, vectors $(\xi_n)_{n \geqslant 1}$ that are orthonormal in $\mathscr{H}'$, and a self-adjoint operator E on $\overline{\mathscr{H}}$ can all be found such that $U = \exp(-i\tau E)$ satisfies the von Neumann criteria for measurement of A for some initial state ξ in $\mathscr{H}'$, then for U to commute with $L = L_1 \otimes 1 + 1 \otimes L_2$ it is necessary that A commutes with L_1. Their work also showed that when $\mathscr{H}$ and $\mathscr{H}'$ are finite dimensional and the cardinality of the spectrum of L_2 is allowed to be sufficiently large, then approximate measurements of A (in a certain sense) can be made with any desired accuracy. This raises the question, when A, ξ, and $\mathscr{H}'$ are fixed, of explicitly calculating the lower bound of the accuracy of approximate measurements of A in terms of the variance of L_2 in the state ξ. For spin measurements of spin $\frac{1}{2}$ particles (when the conserved quantity is the z-component of the angular momentum), M. Yanase (Phys. Rev., **123** (1961), pp. 666–668) has investigated this question.

6. *Quantum Probability Theory.* For a given logic $\mathscr{L}$ the notions of probability measures on $\mathscr{L}$ and observables associated to $\mathscr{L}$ define a non-commutative generalization of conventional probability theory. For standard logics one can develop such a generalization in great depth: see for instance the articles in *Quantum Probability and Applications to the Quantum Theory of Irreversible Processes*, Springer Lecture Notes in Mathematics, **1055** (1984), edited by L. Accardi, L. Frigerio and V. Gorini.

One of the concepts that is needed for such a theory of Quantum Probability, and which goes beyond what we have treated in Chapters III and IV is that of *conditional probability*. Let $\mathscr{H}$ be a complex separable Hilbert space and $\mathscr{L}=\mathscr{L}(\mathscr{H})$. Suppose p is a probability measure on $\mathscr{L}$ and U the associated von Neumann operator. We identify elements of $\mathscr{L}$ with the orthogonal projections corresponding to them. If E is a projection and $p(E)>0$, the *conditional probability measure given that E has occurred* is defined as the probability measure $p\,(\,.\,|E)$ given by

$$p(F|E)=\operatorname{tr}(EUEF)/\operatorname{tr}(EUE). \tag{42}$$

It is not difficult to show that $p(\,.\,|E)$ is the *unique* probability measure on $\mathscr{L}$ with the following property: to any $F\leqslant E$, it assigns the probability $p(F)/p(E)$. Clearly $p(\,.\,|E)$ is defined by the von Neumann operator V where

$$V=(\operatorname{tr}(UE))^{-1}\cdot EUE. \tag{43}$$

If F is a projection commuting with E we have

$$p(F|E)=p(EF)/p(E), \tag{44}$$

which is exactly as in the conventional probability theory; however, for *arbitrary* F we should use (42). If $p=p_\varphi$ where $\varphi\in\mathscr{H}$ and $\|\varphi\|=1$, we have

$$p(\,.\,|E)=p_\psi, \qquad \psi=\|E\varphi\|^{-1}\cdot E\varphi. \tag{45}$$

Let $E=E_1+E_2+\ldots$ be an orthogonal sum of projections $E_1,E_2,\ldots$ with $p(E_n)>0$ for all n. If the projection F commutes with all the E_n, we have

$$p(F|E)=\sum_n\,(p(F|E_n)\,p(E_n)/p(E)). \tag{46}$$

However, for *general* F, (46) *is no longer true*. If $p=p_\varphi$ ($\varphi\in\mathscr{H}$, $\|\varphi\|=1$) and $\psi_n=\|E_n\varphi\|^{-1}\cdot E_n\varphi$, then for an arbitrary projection F we have the correct formula

$$p(F|E)=\sum_n\left(\frac{\|E_n\varphi\|}{\|E\varphi\|}\right)^2 p(F|E_n)+\sum_{m\neq n}\left(\frac{\|E_m\varphi\|\,\|E_n\varphi\|}{\|E\varphi\|^2}\right)(F\psi_m,\psi_n). \tag{47}$$

The first term in the right side of this formula is the same as that in the right side of (46). The additional terms in the right side of (47) may be viewed as arising out of the "interference effects" that are typical in quantum theory when F does not commute with the E_n. For a very interesting discussion of the well-known two-slit experiments of Quantum Mechanics from the point of view of conditional probability see Chapter 26 of the book by E. Beltrametti and G. Cassinelli (cf. Notes on Chapter III). We refer to the same place for a treatment of conditional probabilities relative to a Boolean sub-σ-algebra of $\mathscr{L}$, and, more generally, relative to a sublogic of $\mathscr{L}$.

7. *Measures on the projections of a von Neumann algebra: Generalizations of Gleason's theorem.* Gleason's theorem may be viewed as a complete description of all countably additive probability measures on the standard logic $\mathscr{L}(\mathscr{H})$ when $\dim(\mathscr{H}) \geqslant 3$. It is natural to ask whether similar results can be proved for probability measures on the logic of projections of general von Neumann algebras. This question was first raised by G. W. Mackey (Amer. Math. Monthly, **64** (1957), pp. 45–57), and it appears that it has now been settled completely.

Let $\mathscr{H}$ be a complex separable Hilbert space and $\mathscr{A}$ a von Neumann algebra of bounded operators on $\mathscr{H}$. We write $\mathscr{A}^P$ for the logic of projections in $\mathscr{A}$ and $\mathscr{A}^{sa}$ for the real vector space of all self-adjoint elements of $\mathscr{A}$. By a finitely additive probability measure on $\mathscr{A}^P$ we mean a map $\mu : \mathscr{A}^P \to [0,1]$ such that $\mu(0) = 0$, $\mu(1) = 1$, $\mu(P+Q) = \mu(P) + \mu(Q)$ for P, $Q \in \mathscr{A}^P$ with $PQ = 0$. Given such a μ, one can associate to it the corresponding expectation functional

$$E_\mu : \mathscr{A}^{sa} \to \mathbf{R}$$

in the following way: if $A \in \mathscr{A}^{sa}$ and $P^A(E \to P_E)$ is the spectral measure of A on the σ-algebra $\mathscr{B}$ of Borel subsets of $\mathbf{R}$, then

$$E_\mu(A) = \int_{-\infty}^{\infty} \lambda d\nu(\lambda), \tag{48}$$

where ν is the finitely additive probability measure $E \to \mu(P_E)$ on $\mathscr{B}$. It is easy to see that the correspondence

$$\mu \to E_\mu$$

is bijective from the set of all finitely additive probability measures on $\mathscr{A}^P$ to the set of all functionals

$$E : \mathscr{A}^{sa} \to \mathbf{R}$$

with the following properties:

(49)
(i) E is linear (over $\mathbf{R}$) on any *abelian* subalgebra of $\mathscr{A}^{sa}$;
(ii) $E(A^2) \geqslant 0$ for any $A \in \mathscr{A}^{sa}$;
(iii) $E(1) = 1$.

Such functionals are called *physical states* of $\mathscr{A}$. The concept of a physical state is in principle weaker than that of a *state* of $\mathscr{A}$ which demands linearity in place of (i) while retaining (ii) and (iii). If $\mathscr{A} = B(\mathscr{H})$, Gleason's theorem implies that for any countably additive μ, E_μ is a state of $B(\mathscr{H})$. The problem of generalizing Gleason's theorem on $\mathscr{A}^P$ may then be split into two parts:

(1) to determine for which $\mathscr{A}$ all physical states are states;
(2) if a state of $\mathscr{A}$ is *normal*, i.e., if it is countably additive on mutually orthogonal projections, to obtain representations of it in terms of the usual trace on $\mathscr{A}$ (whenever such a trace exists).

The first problem focuses directly on the issues arising out of the non-commutativity of the algebra of observables, and was recognized from the beginning to have crucial significance from the foundational point of view. Since the linearity of physical states is no longer true when $\dim(\mathscr{H}) = 2$ it is clear that we must assume that $\mathscr{A}$ has no central direct summand of type I_2. Recent work seems to have now shown that for *such an* $\mathscr{A}$, *all its physical states are actually states*. This result, certainly one of the most fundamental and profound ones in the subject, is the culmination of the efforts and contributions of many people spread over several years. In addition to the famous paper of Gleason the following (partial) list of articles may be cited: J. F. Aaarnes, Trans. Amer. Math. Soc. **149**, (1970), pp. 601–625; J. Gunson, Ann. Inst. H. Poincaré, A **17** (1972), pp. 295–311; A. A. Lodkin, Funk. Anal. Priloz., **8** (1974), pp. 54–58; M. S. Matvieicuk, Funk. Anal. Priloz., **15** (1981), pp. 41–53, and Teor. Mat. Fiz., **48** (1981), pp. 261–265; E. Christensen, Comm. Math. Phys. **86** (1982), pp. 529–538; F. J. Yeadon, Bull. London Math. Soc., **15** (1983), pp. 139–145, and Bull. London Math. Soc., **16** (1984), pp. 145–150; and A. Paszkiewicz, *preprint*, to appear in J. Funct. Anal. For a survey of these and other relevant articles and a discussion of the ideas used in the proof we refer to the article of P. Kruszyński: *Extensions of Gleason's theorem*, in *Quantum Probability and Applications to the Quantum Theory of Irreversible Processes*, Springer Lecture Notes in Mathematics, **1055** (1984), pp. 210–227, edited by L. Accardi, A. Frigerio and V. Gorini.

The second problem is really part of the study of normal states, and may be viewed as a type of Radon–Nikodym theorem. Such theorems go back to H. A. Dye (Trans. Amer. Math. Soc., **72** (1952), pp. 243–280), and the reader should consult the literature on von Neumann algebras for definitive formulations of such results.

CHAPTER V

MEASURE THEORY ON G-SPACES

1. BOREL SPACES AND BOREL MAPS

From this point onwards we shall emphasize some of the more sophisticated and specialized aspects of quantum theory. We shall deal only with complex separable Hilbert spaces, referring to them as Hilbert spaces without any qualification.

We pointed out in Chapter III the role played by representations of physical symmetry groups into the groups $\mathrm{Aut}(\mathscr{L})$ and $\mathrm{Aut}(\mathscr{S})$, where $\mathscr{L}$ is the logic and $\mathscr{S}$ is the state space of some quantum mechanical system. If we assume that $\mathscr{L}$ is standard, then the results of Chapter IV describe rather completely the structure of $\mathscr{L}$ and $\mathscr{S}$. It would thus appear quite feasible to study the representations of the groups which are important from the physical point of view. Such a study would shed considerable light on various aspects of quantum theory. Now, the methods used in the theory of representations of groups are very sophisticated and depend heavily on analysis on homogeneous spaces. The object of this chapter is to present the basic mathematical theory of homogeneous spaces and the function spaces associated with them.

The main reference to the theory of measure and integration on groups and homogeneous spaces is the book of A. Weil [1]. This however deals mostly with compact and abelian groups. The general theory which is presented in the following sections is to a very large extent the work of Mackey [2], [3], [4], [5], [6], [7].

We shall begin with the concept of a Borel space. A *Borel structure* on a set X is simply a σ-algebra $\mathscr{B}$ of subsets of X; the pair $(X,\mathscr{B})$ will be referred to as a *Borel space*. Elements of $\mathscr{B}$ will be referred to as the *Borel subsets* of X. We shall use this terminology only when there is no ambiguity about the σ-algebra involved. By the usual abuse of language we shall refer to X itself as a Borel space. If $(X,\mathscr{B})$ and $(Y,\mathscr{C})$ are Borel spaces, and f is a map of X into Y, f is said to be *Borel* if $f^{-1}(\mathscr{C})\subseteq\mathscr{B}$; if f is one-one, maps X onto Y, and if $f^{-1}(\mathscr{C})=\mathscr{B}$, we shall say that f is a *Borel isomorphism*, and speak of X and Y as being isomorphic. When $X=Y$ and f is an isomorphism we shall call f an *automorphism.*

The theory of Borel spaces is a very extensive one. We shall constantly use the main results of this theory in our work. For this, our references are to Halmos [1], Kuratowski [1], [2]. Articles of Mackey [5] and Blackwell [1] are also very close to our general point of view. In the next few paragraphs we shall summarize what is essential for our purposes; proofs of these and other related results may be found in the sources mentioned above; cf. also Varadarajan [3], [4].

A Borel space $(X,\mathscr{B})$ is said to be *separable* if (i) $\{x\} \in \mathscr{B}$ for all $x \in X$ and (ii) there is a countable set $\mathscr{D} \subseteq \mathscr{B}$ such that $\mathscr{B}$ is the smallest σ-algebra of subsets of X containing $\mathscr{D}$; $\mathscr{D}$ is then said to be a *generating* class. A class $\mathscr{D} \subseteq \mathscr{B}$ is said to be *separating* if, for each pair of points of X, there exists a set $A \in \mathscr{D}$ which contains one but not both of them. It is useful to note that in any Borel space $(X,\mathscr{B})$, if $\mathscr{D}$ generates $\mathscr{B}$ and $\{x\} \in \mathscr{B}$ for all $x \in X$, then $\mathscr{D}$ is separating. In fact, if this is not true, there will be a pair of points $x, y \in X$ $(x \neq y)$ such that for each $A \in \mathscr{D}$, either $\{x,y\} \subseteq A$ or $\{x,y\} \cap A = \varnothing$. The set of all such A is a σ-algebra containing $\mathscr{D}$ and hence every set in $\mathscr{B}$ has the above property. But this is impossible as $\{x\}$ does not have the property.

Let X be a topological space. The smallest Borel structure containing all the open subsets of X is called the *natural* Borel structure of X. When X is a metric space, the natural Borel structure coincides with the smallest Borel structure with respect to which all continuous functions on X are Borel.

Suppose X_α $(\alpha \in J)$ are Borel spaces and X any set. For each $\alpha \in J$ let π_α be a map of X into X_α. Then there exists a unique smallest Borel structure on X with respect to which all the maps π_α are Borel. If Y is a Borel space and f a mapping of Y into X, f is Borel if and only if $\pi_\alpha \circ f$ is Borel for each α. If X is the product space of the X_α and π_α the projection of X on X_α, X, equipped with the above mentioned Borel structure, will be called the *product* of the Borel spaces X_α. If each X_α is a topological space and its Borel structure is the natural one, then it is not in general true that the product Borel structure coincides with the natural Borel structure associated with the product topology on X. This is so, however, if J is countable and each X_α has a second countable metrizable topology.

If $(X,\mathscr{B})$ is a Borel space and Y is a subset of X, then the class $\mathscr{B}_Y$ defined by

$$\mathscr{B}_Y = \{B \cap Y : B \in \mathscr{B}\}$$

is a Borel structure for Y. $(Y,\mathscr{B}_Y)$ is said to be the *Borel subspace defined by* Y. The natural injection of Y into X is Borel. Note that Y *itself need not be a* Borel set in X. If X is a topological space with its natural Borel structure, then $\mathscr{B}_Y$ coincides with the natural Borel structure of Y when Y is considered as a topological space with the relative topology.

An important class of Borel spaces may be singled out in the following way. Let T be a complete separable metric space and X a *Borel subset* of T. Then the Borel subspace defined by X is said to be *standard*. The motivation for this terminology rests on the remarkable theorem that two standard Borel spaces are isomorphic if and only if they have the same cardinal number and that a standard Borel space is finite, countable, or has the power of the continuum. Standard Borel spaces are separable. If $(X,\mathscr{B})$ is a standard Borel space, $(Y,\mathscr{C})$ is separable, and if f is any one-one Borel map of X into Y, then (i) the range Y_f of f is a Borel set in Y, (ii) the Borel subspace defined by Y_f is standard, and (iii) f is a Borel isomorphism of X with Y_f. If $(X_i,\mathscr{B}_i)$ $(i=1,2,\cdots)$ is a sequence of standard Borel spaces, then their product is standard. A Borel subspace of a standard Borel space is standard if and only if it is defined by a Borel set. Any uncountable standard Borel space is isomorphic to the unit interval. If X is any locally compact Hausdorff space satisfying the second axiom of countability, the Borel space obtained by equipping X with its natural Borel structure is standard.

A *measure* on a Borel space is any nonnegative set function which is countably additive on its Borel structure; $+\infty$ is an allowed value. A measure p on $(X,\mathscr{B})$ is *finite* if $p(X)<\infty$; is *σ-finite* if $X=\bigcup_n X_n$, where each X_n is a Borel set and for each n, $p(X_n)<\infty$. p is said to be *standard* if there exists a Borel set X_0 such that (i) $p(X-X_0)=0$ and (ii) the Borel subspace defined by X_0 is standard. Let X be a locally compact Hausdorff space satisfying the second axiom of countability. By a *Borel measure* on X we shall understand a measure with the property that the measure of any compact set is finite. Any Borel measure is σ-finite. Let $C_c(X)$ be the linear space of all complex valued continuous functions on X with compact support and let μ be any Borel measure on X. Then

$$\tilde{\mu} : f \to \int_X f d\mu \qquad (f \in C_c(X))$$

is a linear functional on $C_c(X)$ such that $\tilde{\mu}(f)\geq 0$ whenever f is real and ≥ 0. Conversely, if $\lambda(f\to\lambda(f))$ is a linear functional on $C_c(X)$ such that $\lambda(f)\geq 0$ whenever f is real and ≥ 0, there exists a unique Borel measure μ on X such that $\lambda=\tilde{\mu}$. This is the well known Riesz theorem (cf. Halmos [1], pp. 216–249). Any Borel measure on X is *regular*, i.e., for any Borel set $E\subseteq X$,

$$\mu(E) = \sup_{\substack{C\subseteq E \\ C \text{ compact}}} \mu(C). \tag{1}$$

Given a Borel space $(X,\mathscr{B})$ and a σ-finite measure p on it, we call a set $A\subseteq X$ *p-measurable* if there exist Borel sets B_1 and B_2 such that $B_1\subseteq A\subseteq B_2$ and $p(B_2-B_1)=0$. When there is no ambiguity about the measure that is involved, we speak simply of measurable sets. The collection of p-measurable sets is a σ-algebra and hence defines a Borel structure on X, say $\bar{\mathscr{B}}_p$, called the *p-completion* of $\mathscr{B}$. The measure p has a unique extension to $\bar{\mathscr{B}}_p$ as a measure. The functions which are Borel with respect to the Borel structure $\bar{\mathscr{B}}_p$ are called p-measurable. If f is any p-measurable function on X, there exists a Borel function g such that $f=g$ almost everywhere, i.e., $\{x : f(x)\neq g(x)\} \in \bar{\mathscr{B}}_p$ and has p-measure zero.

Let $(X,\mathscr{B})$ be a Borel space, let q be any σ-finite measure, and let p be a measure. We shall say that p is absolutely continuous with respect to q, $p\ll q$ in symbols, if $p(E)=0$ for any Borel set E for which $q(E)=0$. If f is a nonnegative Borel function, the function $p : E \to \int_E f dq$ is a σ-finite measure which is $\ll q$; p is finite if and only if f is an element of $\mathscr{L}^1(q)$. Conversely, if p is a σ-finite measure $\ll q$, there exists a nonnegative Borel function f on X such that $p(E)=\int_E f dq$ for all Borel sets E. f is essentially uniquely determined by p in the sense that if f' is a Borel function and $p(E)=\int_E f' dq$ for all Borel sets E, then $f=f'$ q-almost everywhere. f is called a *Radon-Nikodym derivative of p with respect to q* and is denoted by dp/dq. If p, q, r are σ-finite measures, then $p\ll q$ and $q\ll r$ implies $p\ll r$ and the relation $dp/dr=dp/dq.\ dq/dr$ holds r-almost everywhere. In particular, if $p\ll q$ and $q\ll p$, p and q have the same null sets and $dp/dq\cdot dq/dp=1$ almost everywhere; both the derivatives are then positive almost everywhere. Let p be any σ-finite measure on X, and let us choose disjoint Borel sets $X_1,X_2,\cdots$ such that $X=\bigcup_n X_n$ and $p(X_n)<\infty$ for all n. Let the constants $c_n>0$ be such that $\sum_n c_n p(X_n)<\infty$. Then, the measure q defined by

$$q(E) = \sum_n c_n p(E \cap X_n) \qquad (E \in \mathscr{B})$$

is finite, and p and q are mutually absolutely continuous. The relation of mutual absolute continuity is obviously an equivalence relation in the set of all σ-finite measures on X. The corresponding equivalence classes are called *measure classes* on X. If t is a Borel automorphism of X and p is a σ-finite measure on X, we define the σ-finite measure p^t by

$$p^t(E) = p(t^{-1}(E)) \qquad (E \in \mathscr{B}) \tag{2}$$

p is said to be *invariant* if $p^t=p$; p is said to be *quasi invariant* if p^t and p are mutually absolutely continuous. Clearly p is quasi invariant if and only if the class of null sets of p is invariant under t. A measure class is said to be *t-invariant* if it contains a σ-finite measure quasi invariant with respect to t.

We shall now prove two lemmas of a technical nature. They are of importance in the sequel.

Lemma 5.1 (*Federer-Morse* [1]). *Let X and Y be two mectric spaces and let f be a map of X onto Y. Suppose that there exists a sequence $\{K_n\}$ of compact subsets of X such that* (i) $K_1 \subseteq K_2 \subseteq \cdots$, (ii) $X = \bigcup_n K_n$, *and* (iii) *for each n, the restriction of f to K_n is a continuous map of K_n into Y. Then there exists a Borel set $E \subseteq X$ such that* (i) *f maps E onto Y,* (ii) *for any $y \in Y$, E meets $f^{-1}(\{y\})$ in exactly one point, and* (iii) *$E \cap f^{-1}(f[K_n]) \subseteq K_n$, for all n.*

Proof. We shall first consider the special case when X itself is compact and f is a continuous map of X onto Y. By a theorem of Alexandroff (cf. Kelley [1], Chapter 5), the space X is a continuous image of a compact set $C \subseteq [0,1]$. Let g be a continuous map of C onto X and let $h = f \circ g$. Then h is a continuous map of C onto Y. For any $y \in Y$, $h^{-1}(\{y\})$ is a nonempty compact subset of C and hence has a unique least element $m(y)$, i.e., $m(y) \in h^{-1}(\{y\})$ and $m(y) < c$ for all $c \in h^{-1}(\{y\}) - \{m(y)\}$. Define

$$F = \{m(y) : y \in Y\}$$

and

$$E = g[F].$$

By its definition, F meets each set $h^{-1}(\{y\})$ exactly once. Hence h is one-one and maps F onto Y. From this it follows immediately that f is one-one on E and maps E onto Y. In other words, E meets each set $f^{-1}(\{y\})$ $(y \in Y)$ exactly once.

It remains to prove that E is a Borel set. For any integer $n \geq 1$, let $A_n \subseteq C$ be defined by

$$A_n = \{c : c \in C, h(d) \neq h(c) \text{ for any } d \in C \text{ with } d \leq c - 1/n\}.$$

We claim that A_n is open in C. Suppose this is not true. Then there exists a sequence c_s in C such that $c_s \notin A_n$ for all $s = 1, 2, \cdots$, but $c_s \to c \in A_n$ as $s \to \infty$. Since $c_s \notin A_n$ there exists a $c_s' \leq c_s - (1/n)$ in C such that $h(c_s') = h(c_s)$. If $s_1 < s_2 < \cdots$ is a sequence of integers such that $c_{s_k}' \to c'$ as $k \to \infty$, then $c' \in C$, $c' \leq c - (1/n)$, and $h(c') = h(c)$, a contradiction. A_n is therefore open.

We show next that $F = \bigcap_n A_n$. If $c \in F$, it follows from the definition of $m(y)$ that for any $c' < c$, $h(c') \neq h(c)$. Thus $c \in A_n$ for all n. Conversely, let $c_0 \in A_n$ for all n. If $c' < c_0$, there exists an n such that $c' \leq c_0 - (1/n)$ and hence $h(c') \neq h(c)$. This proves that $c_0 = m(h(c_0))$.

Now any open set in a metric space is a countable union of closed sets. As C is compact, it follows that each A_n is a countable union of compact sets. Therefore $g[A_n]$ also has the same property, and is hence a Borel set. We shall complete the proof that E is Borel by showing that $E = \bigcap_n g[A_n]$. Clearly, $E \subseteq \bigcap_n g[A_n]$. Suppose, on the other hand, that $x \in g[A_n]$ for all n and $x \notin g[F]$. Let $y = f(x)$ and $c = m(y)$. Then $c \in F$ and so $g(c) \neq x$; thus $g^{-1}(\{x\}) = T$ is a compact set $\subseteq C$, not containing c and lying entirely to the right of c. Hence there exists a $d>0$ such that all points of T are at a distance $\geq d$ away from c. If $(1/n)<d$, then $T\subseteq C-A_n$ and as $T=g^{-1}(\{x\})$, $x\notin g[A_n]$. This contradiction shows that $E=\bigcap_n g[A_n]$.

We now come to the general case. Write $X=\bigcup_n K_n$, where K_n is compact, $K_1\subseteq K_2\subseteq\cdots$, and f is continuous on each K_n. There exists, by the special case, proved just now, a Borel set $D_n\subseteq K_n$ such that f is one-one on D_n and $f[D_n]=f[K_n]$. Now, the D_n's need not increase with n. We shall show, however, that they may be chosen to possess this property as well. We define Borel sets $E_1, E_2, \cdots$ by induction in the following manner:

$$E_1 = D_1,$$

and for $n>1$,

$$E_n = \{D_n - D_n \cap (f^{-1}(f[K_{n-1}]))\} \cup E_{n-1}. \tag{3}$$

We claim that the E_n's have all the desired properties. Clearly $E_1\subseteq E_2\subseteq\cdots$. We claim that for each n, f is one-one on E_n and $f[E_n]=f[K_n]$. This is proved by induction on n. Since $E_1=D_1$, this is true for $n=1$. Suppose now that $n>1$ and that f is one-one on E_{n-1} with $f[E_{n-1}]=f[K_{n-1}]$. Then the formula (3) implies easily that f is one-one on E_n and maps E_n onto $f[K_n]$. Let $E=\bigcup_n E_n$. Then E is a Borel set, f is one-one on E, and $f[E]=Y$. Since $E_n\subseteq K_n$, it is obvious that $E\cap f^{-1}(f[K_n])\subseteq K_n$ for all n. The lemma is completely proved.

Corollary 5.2. *Let X and Y be standard Borel spaces and f a Borel map of X into Y. Let p be a finite measure on X and q the measure $E\to p(f^{-1}(E))$ defined on Y. Then, the range $f[X]$ of f is q-measurable and its complement has q-measure zero. Moreover, there exist Borel sets A and Z such that* (i) *$A\subseteq X$, $Z\subseteq f[X]$,* (ii) *$q(Y-Z)=0$, and* (iii) *A is a section for f over Z, i.e., f is one-one on A and maps A onto Z.*

Proof. The corollary is trivial if X is countable. Hence we consider the case when X is uncountable. We may assume that X and Y are both identical with the Borel space associated with the unit interval [0,1]. By a theorem of Lusin (Halmos [1], p. 243) we can find compact sets $K_n\subseteq X$ such that (a) $K_1\subseteq K_2\subseteq\cdots$, (b) $p(X-\bigcup_n K_n)=0$, and (c) f is continuous on each K_n. Let $Z=\bigcup_n f[K_n]$. Each $f[K_n]$ is compact and hence Z is a Borel set. Since $\bigcup_n K_n\subseteq f^{-1}(Z)$, $p(X-f^{-1}(Z))=0$ so that $q(Y-Z)=0$. This already shows that $f[X]$ is q-measurable and $q(Y-f[X])=0$. We can now apply the Federer-Morse lemma to $\bigcup_n K_n$ and construct a Borel set $A\subseteq X$ such that f is one-one on A and maps A onto Z. The proof of the corollary is complete.

Remark. It is actually true (though it is harder to prove) that the range $f[X]$ is a universally measurable set; i.e., for *any* finite measure α on Y, $f[X]$ is an α-measurable set. Corollary 5.2 is valid even in this case, with q replaced by α. We shall not prove these results but refer the reader to the paper of von Neumann [2] and Kuratowski [1], [2]. We refer to this stronger version of corollary 5.2 as *von Neumann's cross-section lemma.*

Suppose X and Y are Borel spaces. Let ν be a measure on Y, and for each $x \in X$ let us be given a ν-equivalence class of Borel functions on Y. The problem we now want to examine is whether we can select for each x a representative $f(x,.)$ in the corresponding class such that the function $x, y \to f(x,y)$ is Borel on $X \times Y$. We need two auxiliary lemmas.

Lemma 5.3. *Let X be a Borel space, S a separable metric space, and f a Borel map of X into S. Then there exists a sequence $\{f_n\}$ of Borel maps of X into S such that* (i) *each f_n takes only countably many values, and* (ii) *$f_n(x) \to f(x)$ uniformly for $x \in X$ (as $n \to \infty$).*

Proof. Fix the integer $n \geq 1$. Since S is a separable metric space, we can cover S by a countable family $\{B_m\}$ of balls of radius $1/2n$. Let the sets C_m be defined by $C_1 = B_1$ and $C_m = B_m - \bigcup_{j<m} B_j$ $(m > 1)$. Then the C_m are disjoint and $S = \bigcup_m C_m$. Since $C_m \subseteq B_m$, diameter of $C_m \leq 1/n$ for all m. For each m such that C_m is nonempty, let s_m be some point of C_m. We define f_n as the map

$$f_n(x) = s_m \quad \text{if} \quad x \in f^{-1}(C_m).$$

Then f_n is Borel, has countably many values at most, and $\operatorname{dist}(f_n(x), f(x)) \leq 1/n$ for all $x \in X$.

Lemma 5.4. *Let X be a Borel space and S a separable Banach space. Let K be a subset of S^*, the dual of S, with the property that the linear combinations of elements of K are dense in S^*. If f is a map from X to S such that for each $k \in K$, the function $x \to k(f(x))$ is Borel on X, then f is a Borel map of X into S.*

Proof. Let $\mathscr{B}^*$ be the smallest Borel structure on S with the property that all the elements of S^* are Borel functions with respect to it. From our assumption about K it follows easily that for each $A \in \mathscr{B}^*$, $f^{-1}(A)$ is a Borel subset of X. To prove that f is Borel, it is enough to prove that $\mathscr{B}^*$ coincides with the natural Borel structure $\mathscr{B}$ of S. Obviously, $\mathscr{B}^* \subseteq \mathscr{B}$. Since $\mathscr{B}$ is generated by the open sets, it suffices to prove that any open set belongs to $\mathscr{B}^*$. Now, as S is separable, every open set is a union of countably many balls. Thus it is enough to prove that $\mathscr{B}^*$ contains all balls. Further, as $x \to cx$ $(c \in \mathbf{C})$ and $x \to x + a$ $(a \in S)$ are easily seen to be automorphisms of the Borel space $(S, \mathscr{B}^*)$, we are reduced to proving that the unit ball $B = \{x : \|x\| \leq 1\}$ lies in $\mathscr{B}^*$. Now, it is well known that under the weak $*$-topology for S^* (= the smallest topology for S^* such that $k \to k(x)$ is a continuous function on S^* for each $x \in S$), the unit ball B^* of

S^* is a compact metric space (Dunford-Schwartz [1]), hence separable. Let $\{x_1^*, x_2^*, \cdots\}$ be a countable dense subset of the unit ball of S^* in the weak $*$-topology. Then (by the Hahn-Banach theorem),

$$\begin{aligned} B &= \{x : |x^*(x)| \leq 1 \text{ for all } x^* \in B^*\} \\ &= \bigcap_n \{x : |x_n^*(x)| \leq 1 \text{ for all } n\}. \end{aligned}$$

This shows that $B \in \mathscr{B}^*$ and completes the proof that $\mathscr{B}^* = \mathscr{B}$. This proves that f is a Borel map.

Now we are in a position to formulate and prove the lemma on the selection of representatives.

Lemma 5.5. *Let X and Y be Borel spaces, with Y separable. Let μ and ν be σ-finite measures on X and Y, respectively. Let f be a complex valued function on $X \times Y$ and let Y be a union of Borel sets Y_i of finite ν-measure such that* (i) *for each $x \in X$, $y \to f(x,y)$ is a Borel function on Y which is ν-integrable on each Y_i,* (ii) *for any i and any Borel set $F \subseteq Y_i$, $x \to \int_F f(x,y)d\nu(y)$ is a Borel function of x. Then, there exists a complex Borel function f^* on $X \times Y$ and a Borel set $N \subseteq X$ of μ-measure zero such that for each $x \in X - N$,*

$$f^*(x,y) = f(x,y) \tag{4}$$

for ν-almost all y.

Proof. We may clearly drop down on one of the Y_i so that we may assume that $\nu(Y) < \infty$. Also since μ intervenes only in terms of its null sets, we may replace μ by a finite measure mutually absolutely continuous with respect to μ; hence we assume that μ is also finite.

Write $S = \mathscr{L}^1(\nu)$, the Banach space of (equivalence classes of) ν-integrable Borel functions on Y. S is a separable Banach space since the Borel structure of Y is countably generated. Since $y \to f(x,y)$ is a ν-integrable Borel function on Y, it defines an element, say $f(x)$, of S. We claim that $x \to f(x)$ is a map of X into S which satisfies the conditions of lemma 5.4. Now the dual of $\mathscr{L}^1(\nu)$ is $\mathscr{L}^\infty(\nu)$ and every bounded Borel function on Y is a uniform limit of linear combinations of characteristic functions χ_F of Borel subsets F of Y. χ_F defines the linear functional $t_F : u \to \int_F u(y)d\nu(y)$ on $\mathscr{L}^1(\nu)$. By hypothesis, $x \to t_F(f(x))$ is Borel for all F. We may therefore conclude from lemma 5.4 that $x \to f(x)$ is a Borel map from X to S. So X is the disjoint union of countably many Borel sets on each of which f is bounded. We may assume therefore that f itself is bounded. By lemma 5.3 there exists a sequence $\{f_n\}$ of Borel maps of X into S such that (i) each f_n is bounded and takes only countably many values, and (ii) $f_n(x) \to f(x)$ uniformly in $x \in X$. Since f_n is bounded and takes only countably many values, it is obvious that there exists a $\mu\times\nu$-integrable Borel function f_n^* on $X\times Y$ such that for each $x\in X$, the function $y\to f_n^*(x,y)$ defines the element $f_n(x)$ of S. We then have:

$$\sup_{x\in X}\int_Y |f_n^*(x,y)-f_m^*(x,y)|d\nu(y)\to 0$$

as n, $m\to\infty$ and hence

$$\iint_{X\times Y} |f_n^*(x,y)-f_m^*(x,y)|d(\mu\times\nu)(x,y)\to 0$$

as n, $m\to\infty$. As $\mathscr{L}^1(\mu\times\nu)$ is complete, there is a Borel function f^* on $X\times Y$ such that f^* is $(\mu\times\nu)$-integrable and

$$\iint_{X\times Y} |f_n^*(x,y)-f^*(x,y)|d(\mu\times\nu)(x,y)\to 0$$

as $n\to\infty$. Select a subsequence $\{f_{n_k}^*\}$ such that $f_{n_k}^*(x,y)\to f^*(x,y)$ for $(\mu\times\nu)$-almost all x, y. Then there exists a Borel set N of μ-measure zero such that for each $x\in X-N$, $f_{n_k}^*(x,y)\to f^*(x,y)$ for ν-almost all $y\in Y$. It is clear that for each $x\in X-N, f^*(x,y)=f(x,y)$ for ν-almost all y. This proves the lemma.

Corollary 5.6. *Let X, Y, μ, and ν satisfy the same restrictions as in lemma* 5.5. *Suppose that $Q(x\to q_x)$ is a map from X into the space of all finite measures on Y such that* (i) *$q_x\ll\nu$ for all $x\in X$, and* (ii) *for each Borel set $F\subseteq Y$, $x\to q_x(F)$ is a Borel function on X. Then, there exists a Borel function f^* on $X\times Y$ and a Borel set $N\subseteq X$ of μ-measure zero such that for each $x\in X-N$, $y\to f^*(x,y)$ is a version of $dq_x/d\nu$.*

Proof. For each $x\in X$ select a Borel function f_x on Y such that $f_x=dq_x/d\nu$, and define f on $X\times Y$ by $f(x,y)=f_x(y)$. The corollary follows from the lemma at once.

2. LOCALLY COMPACT GROUPS. HAAR MEASURE

The groups which are commonly encountered in physical problems, such as the Lorentz group or the rotation group, are topological groups (cf. Pontrjagin [1] for the theory of topological groups). Moreover the underlying topological spaces are locally compact and satisfy the second axiom of countability. Because of this and other technical reasons, we shall restrict ourselves in the sequel to locally compact groups satisfying the second axiom of countability (lcsc). The Borel structures associated with the underlying topologies of these groups are standard. We shall describe in this section a few known facts about lcsc groups and Haar measures. A detailed treatment can be found in Halmos [1] (pp. 266–289).

Suppose that G is a group which is at the same time a Borel space. We shall say that G is a *Borel group* if the map $x, y \to xy^{-1}$ of $G \times G$ into G is Borel; G is called a *separable* (*standard*, etc.) Borel group if the underlying Borel structure is separable (standard, etc.). In this book we shall deal only with separable Borel groups. If G is a separable Borel group, the maps $x \to x^{-1}$, $x \to xh$, and $x \to hx$ ($h \in G$) are Borel automorphisms of the Borel space G. Any lcsc group is a standard Borel group.

Suppose G is any lcsc group. It is a classical fact that there exists a nonzero σ-finite measure μ on G such that μ is *left invariant*, i.e., $\mu(E) = \mu(xE)$ for all Borel sets E and all $x \in G$. μ is essentially unique in the sense that if μ' is another σ-finite left invariant measure, then $\mu' = c \cdot \mu$ for some constant $c > 0$. μ is called a *left Haar measure* on G. $\mu(U) > 0$ for every open set U and μ is a Borel measure. Similarly there exists a nonzero Borel measure μ_r which is *right invariant*, i.e., $\mu_r(E) = \mu_r(Ex)$ for all Borel sets E and $x \in G$. Again, the right invariance determines μ_r up to multiplication by a positive constant. μ_r is called a *right Haar measure* on G.

In general, it is not true that a right Haar measure is left invariant. When this is so we shall say that G is a *unimodular group*. Suppose G is any lcsc group and that μ and μ_r are left and right Haar measures on G. For any $x \in G$, the measure $E \to \mu(Ex)$ is also a left Haar measure and hence there exists a constant $\Delta(x) > 0$ such that

$$\mu(Ex) = \Delta(x)\mu(E) \tag{5}$$

for all E. Clearly Δ does not depend on which left Haar measure we use in (5). The function $\Delta(x \to \Delta(x))$ is known to be a continuous homomorphism of G into the multiplicative group of positive real numbers, i.e.,

$$\begin{aligned} &\text{(i)} \quad \Delta(e) = 1 \qquad (e \text{ the identity of } G), \\ &\text{(ii)} \quad \Delta(xy) = \Delta(x)\Delta(y) \qquad (x, y \in G), \\ &\text{(iii)} \quad \Delta \text{ is continuous.} \end{aligned} \tag{6}$$

(i) and (ii) are easy consequences of (5); the third needs a little work (cf. Loomis [2], pp. 117–120). Δ is called the *modular function* on G. If f is a continuous function vanishing outside a compact set of G, we have, from (5),

$$\int_G f(zx)d\mu(z) = \Delta(x)^{-1} \int_G f(z)d\mu(z). \tag{7}$$

From this it follows that

$$\int_G f(z)\Delta(z)^{-1}d\mu(z) = \int_G f(zx)\Delta(z)^{-1}d\mu(z). \tag{8}$$

The equation (8) shows that the measure μ_r defined by

$$\mu_r(E) = \int_E \Delta(z^{-1})d\mu(z) \tag{9}$$

is a right Haar measure. It follows from (9) that μ_r and μ are mutually absolutely continuous; the Radon-Nikodym derivatives $d\mu_r/d\mu$ and $d\mu/d\mu_r$ are, respectively, the functions $z \to \Delta(z)^{-1}$ and $z \to \Delta(z)$. A simple calculation using (9) shows that

$$\mu_r(xE) = \Delta(x)^{-1}\mu_r(E) \tag{10}$$

for all Borel sets E and all $x \in G$.

It is clear from (9) that G is unimodular if and only if Δ is the trivial homomorphism, i.e., $\Delta(z)=1$ for all $z \in G$. Since $x \to \log \Delta(x)$ $(x \in G)$ is a continuous homomorphism of G into the additive group of real numbers which does not have any nontrivial compact subgroups, it follows that Δ is trivial on any compact subgroup of G. Compact groups are therefore unimodular; in this case μ and μ_r are finite. It is customary, when G is compact, to reserve the term Haar measure to the unique left and right invariant measure giving measure 1 to the whole of G. If G is not compact, μ and μ_r are not finite. Abelian groups are trivially unimodular. Since Δ is a homomorphism, $\Delta(z)=1$ whenever z is of the form $xyx^{-1}y^{-1}$ for $x, y \in G$. Let G' be the smallest closed subgroup of G containing all elements of the form $xyx^{-1}y^{-1}$ $(x, y \in G)$; Δ is then identically 1 on G'. Hence a sufficient condition for G to be unimodular is that $G=G'$.

There exist groups which are not unimodular. Suppose that G is the group of all matrices g,

$$g = \begin{pmatrix} y & x \\ 0 & 1 \end{pmatrix} \qquad (x, y \text{ real}, y > 0).$$

We identify g with the upper half-plane by the mapping $g \to (x,y)$. If

$$\mu(E) = \iint_E \frac{1}{y^2}\,dxdy, \qquad \mu_r(E) = \iint_E \frac{1}{y}\,dxdy,$$

then μ and μ_r are, respectively, left and right Haar measures on G.

3. G-SPACES

Suppose that X is a Borel space and G is a separable Borel group. We shall say that X is a *G-space* if for each $g \in G$, there exists a Borel automorphism $t_g(x \to g\cdot x)$ of X such that

$$\text{(11)} \qquad \begin{aligned} &\text{(i)} \quad t_e \text{ is the identity,} \\ &\text{(ii)} \quad t_{g_1g_2} = t_{g_1}t_{g_2} \qquad (g_1, g_2 \in G). \end{aligned}$$

X is said to be a *Borel G-space* if the map $g, x \to g \cdot x$ of $G \times X$ into X is Borel. We shall say that X is a *standard* Borel G-space if X is a Borel G-space and if the Borel structure of X is standard. For any G-space X and any $x \in X$, the set

$$G_x = \{g : g \in G, \quad g \cdot x = x\} \tag{12}$$

is a subgroup of G. It is called the *stability subgroup* at x. The set

$$G \cdot x = \{g \cdot x : g \in G\} \tag{13}$$

is called the *orbit* of x. If $y = h \cdot x$ $(h \in G)$, then $G_y = h \cdot G_x \cdot h^{-1}$ as is easily verified. Thus the stability subgroups of the points on the orbit of x are conjugate to G_x and any subgroup conjugate to G_x is the stability subgroup of some point on the orbit of x.

If X is a Borel G-space and Y is a subset of X, we shall say that Y is *G-invariant* if $G \cdot x \subseteq Y$ for all $x \in Y$; in this case, the Borel subspace defined by Y becomes in a natural and obvious fashion, a Borel G-space. It is called the *Borel G-subspace defined by Y*. If X_1 and X_2 are two Borel G-spaces, and f a Borel map of X_1 into X_2, then f is called a *G-homomorphism* if for all $x_1 \in X_1$, and all $g \in G$,

$$f(g \cdot x_1) = g \cdot f(x_1). \tag{14}$$

A *G-isomorphism* is a G-homomorphism f with the additional property that f is a Borel isomorphism of X_1 onto X_2. If there exists a G-isomorphism of X_1 onto X_2, we shall call X_1 and X_2 *isomorphic* Borel G-spaces.

Suppose that X is a locally compact Hausdorff space satisfying the second axiom of countability and, further, that G is a lcsc group. Let X be a G-space with the additional property that the map $g, x \to g \cdot x$ of $G \times X$ into X is continuous. Then X is obviously a Borel G-space. We shall refer to it as a *locally compact G-space*. G is said to act *continuously on X*. In this case, for each $g \in G$, $x \to g \cdot x$ is a homeomorphism of X onto itself. Since G is a countable union of compact sets, $G \cdot x$ is also a countable union of compact sets for any $x \in X$. Every orbit in a locally compact G-space is thus a Borel set. The stability subgroups are, moreover, *closed* subgroups of G.

A measure p on a Borel G-space X is said to be *invariant* if $p(g \cdot E) = p(E)$ for all Borel sets $E \subseteq X$ and all $g \in G$. In many problems it is necessary to work with measures which are only quasi invariant. A measure p on X is said to be *quasi invariant* if $p(E) = 0$ if and only if $p(g \cdot E) = 0$ for $g \in G$, i.e., if the action of G on X transforms p-null sets into p-null sets. Notice that if p is quasi invariant and f_1, f_2 are two functions on X which are equal almost everywhere, their transforms $f_1{}^g$ and $f_2{}^g$ also have this property for any $g \in G$, where $f_i{}^g(x) = f_i(g^{-1} \cdot x)$ $(x \in X, i = 1,2)$. For any measure p on X and any $g \in G$, let us define the measure p^g by

$$p^g(E) = p(g^{-1} \cdot E). \tag{15}$$

p is quasi invariant if and only if p^g and p are mutually absolutely continuous for all $g \in G$. Therefore, if p is quasi invariant, so is any measure mutually absolutely continuous with respect to p. If X is a locally compact G-space and p is a Borel measure, then so is p^g.

A simple example of a G-space (G any lcsc group) is obtained when we take $X = G$ and define, for $x \in G$ and $g \in G$, $g \cdot x = L_g(x) = gx$. Obviously $G = G \cdot e$ and the stability subgroups are all trivial. If μ is a left Haar measure, μ is invariant. If μ_r is a right Haar measure, μ_r is not in general invariant; *but it is quasi invariant* since it is mutually absolutely continuous with respect to μ (cf. (9)).

A Borel G-space X is said to be *transitive* or *homogeneous* if $X = G \cdot x$ for some $x \in X$. In this case, $X = G \cdot y$ for each $y \in X$. The above example is a simple example of a transitive G-space. If X is a standard Borel G-space which is transitive and if the stability groups associated with the points of X are all trivial, then we shall say that X is an *affine* space associated with G. If we assume G is standard and choose a point $x \in X$, the map $g \to g \cdot x$ is a Borel isomorphism of G onto X which is a G-isomorphism of the G-space G (G acting on itself by left translations) on the G-space X. But this isomorphism depends on the choice of x.

We shall conclude this section with a theorem due to Varadarajan [2] which imbeds a standard Borel G-space in a *compact metric* G-space, G being any lcsc group.

Theorem 5.7. *Let G be a lcsc group and X be any standard Borel G-space. Then there exists a compact metric G-space Y, and an invariant Borel set $E \subseteq Y$ such that X is G-isomorphic to the Borel G-subspace of Y defined by E.*

Proof. Let μ be any left Haar measure on G. Denote by $\mathscr{L}^1$ the Banach space of (equivalence classes of) complex Borel functions f on G such that

$$\|f\|_1 = \int_G |f(x)| d\mu(x) < \infty. \tag{16}$$

If $h \in G$ and $f^h(x) = f(h^{-1}x)$ for $x \in G$, $f^h \in \mathscr{L}^1$ and $\|f^h\|_1 = \|f\|_1$. It is well known that $f, h \to f^h$ is a continuous map from $\mathscr{L}^1 \times G$ into $\mathscr{L}^1$ (Loomis [2], p. 118). Let $\mathscr{F}$ be the complex vector space of all complex valued bounded Borel functions on X, and for $u \in \mathscr{F}$ let

$$\|u\|_\infty = \sup_{x \in X} |u(x)|. \tag{17}$$

Since the map $g, x \to g^{-1} \cdot x$ of $G \times X$ into X is Borel, the function $g, x \to u(g^{-1} \cdot x)$ is Borel for each $u \in \mathscr{F}$. If $f \in \mathscr{L}^1$, the function $x \to \int_G f(g) u(g^{-1} \cdot x) d\mu(g)$ is Borel on X; we shall denote it by $f * u$. It is obvious that $f * u \in \mathscr{F}$ and $\|f * u\|_\infty \leq \|f\|_1 \cdot \|u\|_\infty$. $u \to f * u$ is a linear map of $\mathscr{F}$ into itself.

For any nonempty set $\mathscr{D} \subseteq \mathscr{F}$ we write $Y(\mathscr{D})$ for the set of all complex valued maps $\mathbf{c}(f,u \to \mathbf{c}(f,u))$ on $\mathscr{L}^1 \times \mathscr{D}$ such that (i) for each $u \in \mathscr{D}$, $f \to \mathbf{c}(f,u)$ is linear on $\mathscr{L}^1$, and (ii) $|\mathbf{c}(f,u)| \leq \|f\|_1 \cdot \|u\|_\infty$ for all $f, u \in \mathscr{L}^1 \times \mathscr{D}$. We equip $Y(\mathscr{D})$ with the smallest topology with respect to which all the maps $\mathbf{c} \to \mathbf{c}(f,u)(f \in \mathscr{L}^1, u \in \mathscr{D}, \mathbf{c} \in Y(\mathscr{D}))$ are continuous. From the defining properties (i) and (ii) it is obvious that $Y(\mathscr{D})$ is a compact Hausdorff space. If $\mathscr{D}$ is countable, it is clear that $Y(\mathscr{D})$ is even a compact *metric* space. We shall now convert $Y(\mathscr{D})$ into a G-space. We define, for $\mathbf{c} \in Y(\mathscr{D})$ and $g \in G$, the element $g \cdot \mathbf{c}$ by

$$g \cdot \mathbf{c}(f,u) = \mathbf{c}(f^{(g^{-1})},u), \tag{18}$$

where, as usual, $f^h(x) = f(h^{-1} \cdot x)$ $(x, h \in G)$. Since $\|f^{g^{-1}}\|_1 = \|f\|_1$, it is clear that $g \cdot \mathbf{c} \in Y(\mathscr{D})$, and a simple calculation shows that $Y(\mathscr{D})$ is a G-space.

We now claim that the map $g, \mathbf{c} \to g \cdot \mathbf{c}$ of $G \times Y(\mathscr{D})$ into $Y(\mathscr{D})$ is continuous. In view of the definition of the topology on $Y(\mathscr{D})$ this reduces to proving that for each $(f,u) \in \mathscr{L}^1 \times \mathscr{D}$, the map $g, \mathbf{c} \to \mathbf{c}(f^{g^{-1}},u)$ from $G \times Y(\mathscr{D})$ into the complex numbers is continuous. However, an easy calculation shows that the inequality

$$|\mathbf{c}(f^g,u) - \mathbf{c}'(f^h,u)| \leq |\mathbf{c}(f^g,u) - \mathbf{c}'(f^g,u)| + \|f^g - f^h\|_1 \cdot \|u\|_\infty$$

is valid for all $\mathbf{c}, \mathbf{c}' \in Y(\mathscr{D})$ and $g, h \in G$, and the right side tends to 0 when $h \to g$ and $\mathbf{c}' \to \mathbf{c}$. G thus acts continuously on $Y(\mathscr{D})$.

We shall now imbed X in $Y(\mathscr{D})$ for a suitably chosen $\mathscr{D}$. Let $A_1, A_2, \cdots$ be a sequence of subsets of X which generate the Borel structure of X, and let u_n be the function which is 1 on A_n and 0 on $X - A_n$. Let $\mathscr{D} = \{u_1, u_2, \cdots\}$ and let $Y = Y(\mathscr{D})$. Since $\mathscr{D}$ is countable, Y is a compact metric space. For any $x \in X$ let us define $\mathbf{c}_x$ by

$$\mathbf{c}_x(f,u) = (f * u)(x) \qquad (f \in \mathscr{L}^1, u \in \mathscr{D}). \tag{19}$$

Obviously $\mathbf{c}_x \in Y$. $x \to \mathbf{c}_x$ is thus a map, say ξ, of X into Y. Since the map $x \to (f * u)(x)$ is Borel for all fixed f, u, it follows that ξ is a Borel map of X into Y. Further, if $g \in G$, we have:

$$\begin{aligned} \mathbf{c}_{g \cdot x}(f,u) &= (f * u)(g \cdot x) \\ &= \int_G f(h) u((h^{-1} \cdot g) \cdot x) d\mu(h) \\ &= \int_G f(gt) u(t^{-1} \cdot x) d\mu(t) \\ &= (f^{(g^{-1})} * u)(x) \\ &= (g \cdot \mathbf{c})(f,u). \end{aligned}$$

This proves that ξ is a Borel G-homomorphism from the G-space X into the G-space Y. Let E be the range of ξ. E is clearly invariant.

We assert that ξ is one-one. Suppose in fact that x, y are two points of X such that $(f * u)(x) = (f * u)(y)$ for all $f, u \in \mathscr{L}^1 \times \mathscr{D}$. Then, for each $u \in \mathscr{D}$, $u(g^{-1}\cdot x) = u(g^{-1}\cdot y)$ for almost all $g \in G$. As $\mathscr{D}$ is countable, there exists a Borel null set $N \subseteq G$ such that for any $g \in G - N$, $u(g^{-1}\cdot x) = u(g^{-1}\cdot y)$ for all $u \in \mathscr{D}$. Let g_0 be any point of $G - N$, and let $x_0 = {g_0}^{-1}\cdot x$, $y_0 = {g_0}^{-1}\cdot y$. As the Borel space X is separable and as $A_1, A_2, \cdots$ is a generating family, it is also a separating family (cf. Section 1). Hence, $u(x_0) = u(y_0)$ for all $u \in \mathscr{D}$ implies that $x_0 = y_0$. Therefore $x = y$.

Therefore ξ is a one-one Borel map of X into Y. Since X and Y are standard, this implies that the range E of ξ is a Borel subset of Y and that ξ is a Borel isomorphism of the Borel space X onto the Borel subspace of Y defined by E. Since we have already proved that E is an invariant subset of Y and that ξ is a G-homomorphism, the proof of theorem 5.7 is complete.

Corollary 5.8. *Let G be a lcsc group and let X be a standard Borel G-space. Then, for any $x \in X$, the stability subgroup G_x (of G) at x is closed. Moreover, the orbit $G\cdot x$ of any x in X is a Borel subset of X.*

In fact, it is enough, in view of theorem 5.7, to prove these facts when X is an invariant Borel subset of a compact metric space on which G acts continuously. In this case, both assertions are obvious.

We shall end this section with two technical results which describe the manner in which the quasi-invariant measures on a Borel G-space transform under the action of G. For any σ-finite measure p on X we define p^g by (15).

Lemma 5.9. *Let X be any Borel G-space, G, a lcsc group. Let p be a σ-finite quasi-invariant measure on X and let $f_p(g,.)$ be a version of the Radon-Nikodym derivative $dp/dp^{(g^{-1})}$. Then for fixed $g_1, g_2 \in G$,*

$$f_p(g_1g_2,x) = f_p(g_1,g_2\cdot x)f_p(g_2,x) \tag{20}$$

for p-almost all x.

Proof. We begin the proof with a formula which describes how a Radon-Nikodym derivative transforms under a Borel automorphism. Let μ, ν be σ-finite measures on X such that $\mu \ll \nu$. Let $t(x \to t\cdot x)$ be a Borel automorphism of X. Define the measures μ^t and ν^t by setting $\mu^t(E) = \mu(t^{-1}(E))$ and $\nu^t(E) = \nu(t^{-1}(E))$, for all Borel sets E. Then $\mu^t \ll \nu^t$. If E is any Borel set,

$$\begin{aligned}\mu^t(E) &= \int_{t^{-1}(E)} (d\mu/d\nu)(x)d\nu(x)\\ &= \int_E (d\mu/d\nu)(t^{-1}\cdot x)d\nu^t(x),\end{aligned}$$

showing that

$$(21) \qquad d\mu^t/d\nu^t = (d\mu/d\nu)^t.$$

This said, we come to the actual proof. Fix $g_1, g_2 \in G$. Then $p^{(g_2g_1)} = (p^{g_1})^{g_2}$ and $dp/dp^{(g_2g_1)} = (dp/dp^{g_2})\cdot(dp/dp^{g_1})^{g_2}$ from what we said above. This implies easily that

$$f_p(g_1g_2,x) = f_p(g_1,g_2\cdot x)f_p(g_2,x)$$

for p-almost all x.

Our next result tells us that we can choose these Radon-Nikodym derivatives so they are "almost" Borel on $G \times X$.

Theorem 5.10. *Let G be a lcsc group, μ a left Haar measure on G, and let X be a separable Borel G-space. Let p be a σ-finite quasi-invariant measure on X. Then, there exists a Borel function F_p on $G \times X$ which is positive everywhere such that for μ-almost all $g \in G$, $x \to F_p(g,x)$ is a version of the Radon-Nikodym derivative $dp/dp^{(g^{-1})}$. In particular, the identity*

$$(22) \qquad F_p(g_1g_2,x) = F_p(g_1,g_2\cdot x)F_p(g_2,x)$$

is satisfied for $\mu \times \mu \times p$-almost all $(g_1,g_2,x) \in G \times G \times X$.

Proof. We start with the special case when p is finite. We first claim that for each Borel set $E \subseteq X$, $g \to p(g\cdot E)$ is Borel. In fact, the map $t : g, x \to g, g\cdot x$ is a Borel automorphism of $G \times X$ and hence $A = t[G \times E]$ is a Borel subset of $G \times X$. Let μ be a left Haar measure on G. Then we know that for each $g \in G$, the set $A^g = \{x : (g,x) \in A\}$ is a Borel subset of X and $g \to p(A^g)$ is a Borel function on G. Since $A^g = g\cdot E$, our claim is proved.

We are now in a position to apply corollary 5.6. It follows from this corollary that there exists a Borel function F_1 on $G \times X$ such that for almost all $g \in G$, $x \to F_1(g,x)$ is a version of $dp^{(g^{-1})}/dp$. Since p is quasi invariant, for each g, $F_1(g,x)$ is p-almost everywhere positive and hence, by the Fubini theorem, F_1 is positive for almost all points of $G \times X$. We may thus assume that F_1 is positive everywhere without altering the conclusions. Put $F_p = 1/F_1$. Then F is a positive Borel function on $G \times X$ and for μ-almost all g, $x \to F_p(g,x)$ is a version of $dp/dp^{(g^{-1})}$. Finally, the identities (22) follow from (20). This completes the proof of the theorem when p is finite.

If p were σ-finite, we choose a finite measure q such that $q \ll p$ and $p \ll q$. Let $a(x \to a(x))$ be a positive Borel function which is a version of the Radon-Nikodym derivative dp/dq. Choose F_q as a Borel function on $G \times X$ corresponding to q, and define

$$(23) \qquad F_p(g,x) = a(x)\cdot F_q(g,x)\cdot a^{-1}(g\cdot x).$$

The identity

$$dp/dp^{(h)} = dp/dq \cdot dq/dq^{(h)} \cdot (dq/dp)^{(h)}$$

implies that F_p has all the required properties.

4. TRANSITIVE G-SPACES

We shall now examine the transitive G-spaces somewhat more closely. Let G be a lcsc group. Let G_0 be a closed subgroup of G. Let $X = G/G_0$, the space of all left cosets gG_0, $g \in G$. Let β be the canonical mapping of G onto X defined by

$$\beta : g \to gG_0, \tag{24}$$

and for any element g of G and $x = hG_0$ of X let $g \cdot x$ be defined by

$$g \cdot x = ghG_0. \tag{25}$$

Clearly, $g \cdot x$ depends only on g and x and not on the element h used to define x. G thus acts on X and this action is transitive. Let X be equipped with the quotient topology, that is, $U \subseteq X$ is open if and only if $\beta^{-1}(U)$ is open in G. A map $c(x \to c(x))$ of X into G is called a *Borel (continuous) section* if it is a Borel (continuous) map and if $\beta(c(x)) = x$ for all $x \in X$; then its range meets each left G_0 coset exactly once. It is said to be *regular* if for each compact subset K of G, $c[X] \cap \beta^{-1}(\beta[K])$ is a set with compact closure.

Theorem 5.11. *X, under the quotient topology, is a locally compact Hausdorff space satisfying the second axiom of countability and β is an open continuous map. Moreover, X is a G-space and the map $g, x \to g \cdot x$ of $G \times X$ into X is continuous. If K is any compact subset of X, there exists a compact subset K_1 of G such that $\beta[K_1] = K$. Further, there exists a regular Borel section for G/G_0. A map f of X into some Borel space Z is Borel if and only if the map $f^0 = f \circ \beta$ of G to Z is Borel. In particular, if $E \subseteq X$, E is a Borel set if and only if the set $\beta^{-1}(E)$ is a Borel subset of G. Finally, let Y be a transitive standard Borel G-space, $y \in Y$ and let G_0 be the stability subgroup at y. Then the map $t : \beta(g) \to g \cdot y$ is a Borel isomorphism of $X = G/G_0$ onto Y which is a G-isomorphism of the associated G-spaces. If Y is a locally compact G-space satisfying the second axiom of countability, t is a homeomorphism.*

Proof. That X is locally compact Hausdorff second countable and that β is open are well known (cf. Pontrjagin [1]).

Now, G is clearly transitive on X. The map $(g,g') \to (g,\beta(g'))$ of $G \times G$ onto $G \times X$ is continuous and it is easily shown that the topology for $G \times X$ is the quotient topology relative to this map. The map $g, x \to g \cdot x$ is therefore continuous.

Let $W_1, W_2, \cdots$ be a sequence of compact sets in G with nonempty interiors such that $\bigcup_n (\text{interior}(W_n)) = G$. If then K is a compact set $\subseteq X$, then for some n,

$$K \subseteq \bigcup_{j=1}^{n} \beta[W_j],$$

and hence $K = \beta[K_1]$, where $K_1 = \beta^{-1}(K) \cap (\bigcup_{j=1}^{n} W_j)$. Clearly K_1 is compact.

We shall next prove the existence of a regular Borel section for G/G_0. Let $\{K_n\}$ be a sequence of compact subsets of G such that (a) $K_1 \subseteq K_2 \subseteq \cdots$ and $\bigcup_n K_n = G$, and (b) any compact subset of G is contained in some K_n. Using lemma 5.1 we can construct a Borel set E such that (i) $E \cap \beta^{-1}(\beta[K_n]) \subseteq K_n$ for all n, and (ii) β is one-one on E and maps E onto X. β is thus a Borel isomorphism of E onto X as E and X are both standard. If c denotes the inverse map, c is a Borel isomorphism of X onto E. c is a Borel section. It is obvious from (i) and (b) above that c is regular.

Suppose now the set $A \subseteq X$ is such that $\beta^{-1}(A)$ is a Borel set in G. Then $E \cap \beta^{-1}(A)$ is also Borel. Since β is one-one on E, and since the Borel subspace defined by E is standard, $A = \beta[E \cap \beta^{-1}(A)]$ is a Borel subset of X. If f is a Borel map of X into Z, where Z is some Borel space, then $f^0 = f \circ \beta$ is a Borel map of G into Z. Conversely, let f^0 be a Borel map of G into Z such that $f^0 = f \circ \beta$ for some map f of X into Z. If F is any Borel subset of Z, $\beta^{-1}(f^{-1}(F)) = f^{0-1}(F)$ is Borel in G and hence $f^{-1}(F)$ is Borel in X. f is thus a Borel map.

We finally come to the uniqueness of transitive G-spaces. Let Y, y be as in the statement of the last assertion of theorem 5.11. Clearly the map

$$t : \beta(g) \to g \cdot y$$

is one-one, maps X onto Y, and is moreover such that $t(g \cdot x) = g \cdot t(x)$ for all $x \in X$ and $g \in G$. If $M \subseteq Y$ is open, $t^{-1}(M) = \beta[A]$, where $A = \{g : g \cdot y \in M\}$ so that $t^{-1}(M)$ is open. t is thus continuous. A classical category argument (Pontrjagin [1]) now implies that t is a homeomorphism.

Suppose we assume in the above proof that Y is only a standard Borel G-space. We then use corollary 5.8 to conclude that G_y is a closed subgroup of G. Thus t is well defined and the same argument as that given above now shows that t is Borel. Since both X and Y are standard, t is a Borel isomorphism. This completes the proof of the theorem.

Since G acts on X, we may speak of invariant measure classes on X. We shall now proceed to study these. We use a "lifting" technique to reduce the problem to G itself. Now G acts on itself by left and right translations. We have accordingly left and right invariant measure classes. We write μ for a left Haar measure and define μ_r by

$$\mu_r(E) = \mu(E^{-1}). \tag{26}$$

μ_r is a right Haar measure.

Lemma 5.12. *There exists a unique measure class on G which is left (or right) invariant, and it is the measure class determined by μ. In particular, it is both left and right invariant.*

Proof. Let α be a finite right quasi-invariant measure. We want to prove that $\alpha(E)=0$ if and only if $\mu(E)=0$. Let f be a Borel function on G such that

$$\text{(i)} \quad 0 < f(x) < 1 \qquad \text{for all } x \in G,$$

$$\text{(ii)} \quad \int_G f(x)d\mu(x) + \int_G f(x)d\mu_r(x) < \infty.$$

Such f are easily constructed. Suppose now for a Borel set A, $\alpha(A)=0$. Since α is right quasi invariant, $\alpha(Ax)=0$ for all x in G and hence

$$\int_G \alpha(Ax^{-1})f(x)d\mu(x) = 0,$$

i.e., writing χ_A for the characteristic function of A,

$$\int_G \left(\int_G \chi_A(yx)d\alpha(y)\right)f(x)d\mu(x) = 0.$$

By Fubini's theorem, this means

$$\int_G \left(\int_G \chi_A(yx)f(x)d\mu(x)\right)d\alpha(y) = 0$$

and hence, as the inner integral is ≥ 0,

$$\int_G \chi_A(yx)f(x)d\mu(x) = 0$$

for α-almost all y. Thus, for some $y_0 \in G$,

$$\int_G \chi_A(y_0x)f(x)d\mu(x) = 0$$

and hence, as $f(x) > 0$ for all x,

$$\chi_A(y_0x) = 0$$

for μ-almost all x, i.e., $\mu(y_0{}^{-1}A)=0$ so that $\mu(A)=0$. Conversely, if $\mu(A)=0$, $\mu(y^{-1}A)=0$ for each y and hence

$$\int_G \chi_A(yx)f(x)d\mu(x) = 0$$

for each y, proving that

$$\int_G \left(\int_G \chi_A(yx)f(x)d\mu(x)\right)d\alpha(y) = 0.$$

Using Fubini's theorem, we get:

$$\int_G \alpha(Ax^{-1})f(x)d\mu(x) = 0$$

and hence, as the integrand is ≥ 0,

$$\alpha(Ax_0^{-1}) = 0$$

for some x_0, showing, as α is right quasi invariant, that $\alpha(A) = 0$. Thus μ_r, α, μ are all in the same measure class. If α is left quasi invariant, the measure $\hat{\alpha}$ defined by $\hat{\alpha}(A) = \alpha(A^{-1})$ is right quasi invariant so that μ, α, μ_r are in the same measure class. This proves the lemma.

Remark. Lemma 5.12 enables us to speak without ambiguity of null sets in G; a Borel set $A \subseteq G$ will be called null if it is a set of measure zero for the unique measure class singled out by lemma 5.12.

We now return to the general transitive case. Let G_0 be a closed subgroup of G and $X = G/G_0$. We now observe that, on X, quasi-invariant measures can be constructed easily. Let α_0 be a finite measure on G which is quasi invariant, and let α be the measure on X defined by $\alpha(A) = \alpha_0(\beta^{-1}(A))$. Then $\alpha(E) = 0$ if and only if $\beta^{-1}(E)$ is a null set in G. α is thus quasi invariant. Our aim is to show that all quasi-invariant measures can be obtained in this manner, up to absolute continuity.

Let $C_c(G)$ denote the class of all complex functions on G which are continuous and have compact support. Let $C_c(X)$ denote the analogously defined space of functions on X. For any f in $C_c(G)$ let us write Mf for the function on G defined by

$$(Mf)(g) = \int_{G_0} f(gh)d\mu_0(h); \tag{27}$$

here μ_0 is a left Haar measure on the lcsc group G_0 chosen in some way. It is then clear from the left invariance of μ_0 that, for any h_0 in G_0,

$$(Mf)(gh_0) = (Mf)(g) \tag{28}$$

for all $g \in G$. Hence there exists a function $f^{\sim}$ on X such that

$$f^{\sim}(\beta(g)) = (Mf)(g) \qquad (g \in G). \tag{29}$$

We shall use the mapping $f \to f^{\sim}$ to lift measures from X to G.

Lemma 5.13. *Let P be the set of all real nonnegative elements of $C_c(X)$ and P' a subset of P having the following properties:* (i) *if $f_1, f_2 \in P'$ and c_1, c_2 are constants ≥ 0, $c_1f_1 + c_2f_2 \in P'$,* (ii) *if $f \in P'$ and $f_1 \in P$, $f_1f \in P'$, and* (iii) *for any $x \in X$, there is an $f \in P'$ with $f(x) > 0$. Then $P' = P$.*

Proof. Let $x \in X$. Then there exists an $f_x \in P'$ such that $f_x(x) = 1$. Therefore, for some open set U_x with compact closure and containing x, $f_x(y) \geq \frac{1}{2}$ for all $y \in U_x$. Let f_1 be an element of P such that $f_1(y) = f_x(y)^{-1}$ for all $y \in U_x$. Then $f_x' = f_1f_x \in P'$ and has the property that $f_x'(y) = 1$ for all $y \in U_x$. Suppose now that K is any compact subset of X. For each

$x \in K$, let U_x be an open set containing x and f_x' an element of P' such that $f_x' = 1$ on U_x. Choose points $x_1, \cdots, x_k \in K$ such that $K \subseteq \bigcup_{j=1}^{k} U_{x_j}$. Then $f' = f'_{x_1} + \cdots + f'_{x_k}$ belongs to P' and $f'(y) \geq 1$ for all $y \in K$. Choose an $f_2 \in P$ such that $f_2(y) = f'(y)^{-1}$ for all $y \in K$. Then $f'' = f_2 f'$ belongs to P' and $f''(y) = 1$ for all $y \in K$. Finally, let $f \in P$ and let K be a compact set such that $f(y) = 0$ for $y \notin K$. Choose an $f'' \in P'$ such that $f''(y) = 1$ for all $y \in K$. Then $f = ff'' \in P'$. This proves the lemma.

We are now ready to formulate the main properties of the map $f \to f^{\sim}$.

Lemma 5.14. *For any* $f \in C_c(G)$, $f^{\sim} \in C_c(X)$. *The map* $f \to f^{\sim}$ *has the following properties:* (a) $f \geq 0$ *implies* $f^{\sim} \geq 0$; (b) $f \to f^{\sim}$ *maps* $C_c(G)$ *onto* $C_c(X)$ *and maps the set of real nonnegative elements of* $C_c(G)$ *onto the set of real nonnegative elements of* $C_c(X)$; (c) *for any continuous function* u *on* X, $(u^0 f)^{\sim} = u f^{\sim}$ (*where* $u^0 = u \circ \beta$) (d) *if* $a \in G$, $(f^a)^{\sim} = (f^{\sim})^a$.

Proof. We first prove that for any f in $C_c(G)$, $f^{\sim} \in C_c(X)$. Clearly if f vanishes outside a compact set $K \subseteq G$, $f^{\sim}$ vanishes outside $\beta[K]$. Therefore there remains only the continuity of $f^{\sim}$. For this it suffices to prove that Mf is continuous on G since then the continuity of $f^{\sim}$ would follow from the fact that X has the quotient topology. Let $g_0 \in G$ and W_0 be a compact neighborhood of e. Now

$$|(Mf)(g) - (Mf)(g_0)| \leq \int_{G_0} |f(gh) - f(g_0 h)| d\mu_0(h).$$

Since $f \in C_c(G)$, given $\varepsilon > 0$, there exists a compact neighborhood $W \subseteq W_0$ around e such that $|f(a) - f(a')| \leq \varepsilon$ whenever $aa'^{-1} \in W$. If f vanishes outside the compact K, and $gg_0^{-1} \in W$,

$$\begin{aligned} |(Mf)(g) - (Mf)(g_0)| &\leq \int_{(g_0^{-1} W^{-1} K) \cap G_0} |f(gh) - f(g_0 h)| d\mu_0(h) \\ &\leq \varepsilon \int_{K' \cap G_0} d\mu_0(h), \end{aligned}$$

where $K' = (g_0^{-1} W_0^{-1} K) \cap G_0$. This proves that Mf is continuous at g_0, finishing the proof that $f^{\sim} \in C_c(X)$.

The map $f \to f^{\sim}$ is linear and it is obvious that $f \geq 0 \Rightarrow f^{\sim} \geq 0$. Next, let u be a continuous function on X and $u^0 = u \circ \beta$. Then

$$M(u^0 f)(g) = u^0(g) \int_{G_0} f(gh) d\mu_0(h).$$

This shows that $M(u^0 f) = u^0 Mf$. Hence $(u^0 f)^{\sim} = u f^{\sim}$. Third, let $a \in G$. Then $(Mf^a)(g) = (Mf)(a^{-1} g)$ from which we conclude that $(f^a)^{\sim} = (f^{\sim})^a$. This proves that (a), (c), and (d) are satisfied. It only remains to prove (b).

Let

$$P' = \{f^{\sim} : f \in C_c(G),\ f \text{ real and } \geq 0\},$$

and let P be the set of all real nonnegative elements of $C_c(X)$. We claim that P' satisfies the conditions of the lemma 5.13. (i) is obvious and (ii) follows from (c). To check that (iii) is satisfied at each point of X, it is enough to do so at $x_0 = \beta(e)$ because G acts transitively on X and P' satisfies (d). But then, if $f \in C_c(G) \geq 0$ and is 1 in a neighborhood of e, $(Mf)(e) = f^{\sim}(x_0) > 0$ obviously. This shows now that $P' = P$. In particular, $C_c(G)$ is mapped onto $C_c(X)$. The lemma is completely proved.

For any Borel measure α on X, the map $f \to \int_X f^{\sim} d\alpha$ is, by virtue of lemma 5.14, a nonnegative linear functional on $C_c(G)$. Let us denote by α^0 the Borel measure defined on G by this linear functional. We thus have a mapping $\alpha \to \alpha^0$ from the set of Borel measures on X into the set of Borel measures on G. Since $f \to f^{\sim}$ has the whole of $C_c(X)$ for its range, it is obvious that ${\alpha_1}^0 = {\alpha_2}^0$ implies $\alpha_1 = \alpha_2$. $\alpha \to \alpha^0$ is therefore a one-one map.

Let us now consider the equation

$$\int_G f d\alpha^0 = \int_X f^{\sim} d\alpha \qquad (f \in C_c(G)). \tag{30}$$

Our next aim is to extend (30) to certain Borel functions on G.

Lemma 5.15. *Let f be a Borel function on G such that*

$$\int_{G_0} |f(gh)| d\mu_0(h) < \infty \tag{31}$$

for each $g \in G$. Then there exists a unique Borel function $f^{\sim}$ on X such that for all $g \in G$,

$$f^{\sim}(\beta(g)) = \int_{G_0} f(gh) d\mu_0(h).$$

If $f \geq 0$, then $f^{\sim} \geq 0$ and we have

$$\int_G f d\alpha^0 = \int_X f^{\sim} d\alpha \tag{32}$$

in the sense that either both sides are infinite or both sides are finite and equal. If f is bounded and vanishes outside a compact subset of G, then $f^{\sim}$ is bounded and vanishes outside a compact subset of X.

Proof. Since $g, h \to f(gh)$ is Borel on $G \times G_0$, it is clear that $g \to \int_{G_0} f(gh) d\mu_0(h)$ is Borel on G. Moreover, this function of g is constant on the left G_0-cosets. Hence, by theorem 5.11, the existence and uniqueness of $f^{\sim}$ are immediate. Also, $f \geq 0$ implies $f^{\sim} \geq 0$. Suppose $f(g) = 0$ for all $g \notin K$,

K being a compact set, and let $|f(g)| \leq A$ for all $g \in G$. Clearly, $f^{\sim}(x)=0$ for $x \notin \beta[K]$. Further, if $x \in \beta[K]$ and we write $x=\beta(g)$ $(g \in K)$,

$$\begin{aligned} |f^{\sim}(x)| &= \left| \int_{G_0} f(gh) d\mu_0(h) \right| \\ &\leq A \int_{g^{-1}K \cap G_0} d\mu_0(h) \\ &\leq A\mu_0(K^{-1}K \cap G_0). \end{aligned}$$

This proves that $f^{\sim}$ is bounded.

We now come to the proof of (32). We consider first the case when $f=\chi_K$, the characteristic function of the compact set $K \subseteq G$. We can choose a sequence $\{f_n\}$ of nonnegative elements in $C_c(G)$ such that $f_n \downarrow \chi_K$ pointwise on G. Then $f_n^{\sim} \downarrow f^{\sim}$ pointwise on X and

$$\begin{aligned} \int_G f d\alpha^0 &= \lim_{n\to\infty} \int_G f_n d\alpha^0 \\ &= \lim_{n\to\infty} \int_X f_n^{\sim} d\alpha \\ &= \int_X f^{\sim} d\alpha. \end{aligned}$$

This proves (32) in this case. Next, for any Borel set $A \subseteq K$, $f=\chi_A$ is bounded and has compact support and hence $\chi_A^{\sim}$ is bounded and has compact support. It is therefore α-integrable and the map $A \rightarrow \int_X \chi_A^{\sim} d\alpha$ is easily verified to be a measure on the Borel sets of K. By what we have seen earlier, this measure coincides with α^0 on all *compact* $A \subseteq K$. Therefore it must coincide with α^0 for all Borel sets A. But then this proves (32) for $f=\chi_A$ $(A \subseteq K)$. At this stage we know that (32) is true for all simple functions f vanishing outside compact sets.

For the general case, let $f \geq 0$ and let $\{f_n\}$ be a sequence of nonnegative simple functions with compact supports such that $f_n \uparrow f$ pointwise on G. Then $f_n^{\sim} \uparrow f^{\sim}$ pointwise on X. Hence by the Fatou-Lebesgue lemma,

$$\begin{aligned} \int_G f d\alpha^0 &= \lim_{n\to\infty} \int_G f_n d\alpha^0 \\ &= \lim_{n\to\infty} \int_X f_n^{\sim} d\alpha \\ &= \int_X f^{\sim} d\alpha. \end{aligned}$$

This completes the proof of the lemma.

Corollary 5.16. *For a Borel set* $A \subseteq X$, $\alpha(A)=0$, *if and only if* $\alpha^0(\beta^{-1}(A))=0$.

Proof. Suppose that A is a Borel set $\subseteq X$ and $\alpha(A)=0$. If K is a compact set $\subseteq \beta^{-1}(A)$, then, $\chi_K{}^{\sim}=0$ outside A and so, by (32),

$$\begin{aligned}\alpha^0(K) &= \int_X \chi_K{}^{\sim} d\alpha \\ &= 0.\end{aligned}$$

This proves, by regularity of α^0, that $\alpha^0(\beta^{-1}(A))=0$. Conversely, let $\alpha^0(\beta^{-1}(A))=0$. Let $K \subseteq A$ be compact. Since $\beta^{-1}(K)$ is closed, there exists an increasing sequence $\{K_n\}$ of compact subsets of $\beta^{-1}(K)$ whose union is $\beta^{-1}(K)$. Then $\alpha^0(K_n)=0$ for all n, so that we can infer from (32) and the nonnegativity of $\chi_{K_n}{}^{\sim}$ that $\chi_{K_n}{}^{\sim}=0$ almost everywhere for each n. Suppose now that $\alpha(K)>0$. Then, there will exist an $x_1 \in K$ such that $\chi_{K_n}{}^{\sim}(x_1)=0$ for all n. Choose a $g_1 \in \beta^{-1}(K)$ such that $\beta(g_1)=x_1$. Then $\mu_0((g_1{}^{-1}K_n) \cap G_0)=0$ for all n. This is absurd as $(g_1{}^{-1}K_n) \cap G_0 \uparrow G_0$. Thus $\alpha(K)$ must vanish. By regularity of α, $\alpha(A)=0$.

Corollary 5.17. *If* α_1 *and* α_2 *are Borel measures on* X, $\alpha_2 \ll \alpha_1$ *if and only if* $\alpha_2{}^0 \ll \alpha_1{}^0$. *In this case,*

$$d\alpha_2{}^0/d\alpha_1{}^0 = (d\alpha_2/d\alpha_1)^0. \tag{33}$$

Proof. Let $\alpha_2 \ll \alpha_1$ and let $t=d\alpha_2/d\alpha_1$. We may assume $t \geq 0$. For any nonnegative $u \in C_c(G)$, the function ut^0 satisfies the conditions of lemma 8.15 and, moreover,

$$(ut^0)^{\sim} = u^{\sim}t.$$

Since $u^{\sim}t \in \mathscr{L}^1(\alpha_1)$, $ut^0 \in \mathscr{L}^1(\alpha_1{}^0)$ and we have:

$$\begin{aligned}\int_G ut^0 d\alpha_1{}^0 &= \int_X u^{\sim} t d\alpha_1 \\ &= \int_X u^{\sim} d\alpha_2 \\ &= \int_G u d\alpha_2{}^0.\end{aligned}$$

This shows at once that $\alpha_2{}^0 \ll \alpha_1{}^0$ and $d\alpha_2{}^0/d\alpha_1{}^0 = t^0$. On the other hand, if $\alpha_2{}^0 \ll \alpha_1{}^0$, we obtain the relation $\alpha_2 \ll \alpha_1$ at once from the previous corollary.

Corollary 5.18. *If* G *is compact, then for any Borel set* $A \subseteq X$, $\alpha(A) = \alpha^0(\beta^{-1}(A))$.

Proof. In this case, on normalizing μ and μ_0 as usual, we see that for any $u \in C_c(X)$, $(u^0)^{\sim}=u$. Hence, $\int_X u d\alpha = \int_G u^0 d\alpha^0$. This implies the corollary.

From lemmas 5.14 and 5.15 the following theorem follows immediately. It is a basic theorem in the theory of quasi-invariant measures on X.

Theorem 5.19. *There are quasi-invariant Borel measures on* $X=G/G_0$, *where* G *is a lcsc group and* G_0 *a closed subgroup. If* λ *is a finite quasi-invariant measure on* G, *and* $\lambda^{\sim}$ *is the finite measure defined on* X *by the equation* $\lambda^{\sim}(A)=\lambda(\beta^{-1}(A))$, *then* $\lambda^{\sim}$ *is quasi invariant. Any two quasi-invariant* σ*-finite measures on* X *are mutually absolutely continuous. If* $E\subseteq X$ *is any Borel set, the quasi-invariant* σ*-finite measures on* X *vanish for* E *if and only if* $\beta^{-1}(E)$ *is a null set of* G. *If* α *is any Borel measure on* X, α *is quasi invariant (invariant) if and only if* α^0 *is a quasi-invariant (left invariant) measure on* G.

Proof. If $g\in G$ and if α is a Borel measure on X, $(\alpha^g)^0=(\alpha^0)^g$ as is easily seen from (d) of lemma 5.14. Moreover, any σ-finite measure is mutually absolutely continuous with respect to some finite measure. Lemmas 5.14 and 5.15 and corollaries 5.16 and 5.17 then lead quickly to the proofs of all the assertions of this theorem.

Theorem 5.19 enables one to speak of null sets of X and G without specifying any quasi-invariant measure. If λ is any quasi-invariant measure on X, the class of λ-measurable sets is independent of λ. In other words, we may also speak of measurable sets and measurable functions on X without ambiguity.

Corollary 5.20. *Let* $A\subseteq X$. *Then* A *is measurable if and only if* $\beta^{-1}(A)$ *is measurable. If* f *is a map of* X *into some Borel space* Z, f *is measurable if and only if* $f^0=f\circ\beta$ *is a measurable map of* G *into* Z.

Proof. It is enough to prove the first part since the second part follows immediately from the first. Suppose now that $A\subseteq X$ is measurable. Then there are Borel sets A_1 and A_2 such that $A_1\subseteq A\subseteq A_2$ and A_2-A_1 is a null set. Clearly $\beta^{-1}(A_1)\subseteq\beta^{-1}(A)\subseteq\beta^{-1}(A_2)$ and $\beta^{-1}(A_2)-\beta^{-1}(A_1)=\beta^{-1}(A_2-A_1)$ is also null. $\beta^{-1}(A)$ is thus measurable.

Conversely, let $B=\beta^{-1}(A)$ be measurable. Let λ be a finite quasi-invariant measure on G and let $\lambda^{\sim}$ be the induced measure on X. There exist sequences $\{K_n\}$ and $\{K_n'\}$ of compact subsets of G such that (i) $K_n\subseteq B$ and $K_n'\subseteq G-B$ for all n, and (ii) $\lambda(B-\bigcup_n K_n)=\lambda((G-B)-\bigcup_n K_n')=0$. Since β is continuous, the sets A_1 and A_2, where

$$A_1=\bigcup_n \beta[K_n]$$

$$X-A_2=\bigcup_n \beta[K_n']$$

are both Borel. It is obvious that $A_1\subseteq A\subseteq A_2$ and $\lambda^{\sim}(A_2-A_1)=0$. A is thus measurable.

Corollary 5.21. *G acts ergodically on X, i.e., if f is a Borel map of X into some Borel space Z such that $f(g\cdot x)=f(x)$ for almost all $(g,x)\in G\times X$, then f is a constant almost everywhere. In particular, there exists, up to a constant multiplying factor, at most one invariant σ-finite measure on X.*

Proof. By the Fubini theorem, there exists a point x_1 of X such that $f(g\cdot x_1)=f(x_1)$ for almost all g. We may clearly assume that $x_1=x_0=\beta(e)$. Select a Borel set $A\subseteq G$ such that A is the union of countably many compact sets, $G-A$ is of Haar measure zero, and $f(g\cdot x_0)=f(x_0)$ for $g\in A$. Let $X_0=A\cdot x_0$. Then X_0 is a Borel set and $\beta^{-1}(X-X_0)\subseteq G-A$ is null, showing that $X-X_0$ is null. Let α_1 and α_2 be nonzero invariant σ-finite measures. Then they are quasi invariant. Let $f=d\alpha_1/d\alpha_2$. Then, for each $g\in G$, $f(g\cdot x)=f(x)$ for almost all x. f is thus a constant almost everywhere, proving that α_1 is a constant multiple of α_2.

Corollary 5.22. *Let S be a standard Borel space and let f be a Borel map of G into S. Suppose that for each $h\in G_0$ $f(gh)=f(g)$ for almost all g. Then there exists a Borel map a of X into S such that $f(g)=a(\beta(g))$ for almost all g.*

Proof. We may clearly assume that S is the unit interval. Let λ be a quasi-invariant measure on G with $\lambda(G)=1$. Let λ_0 be a quasi-invariant measure on G_0 with $\lambda_0(G_0)=1$. Let $\lambda^\sim$ be the measure on X induced by λ through the map β. By the Fubini theorem, $f(gh)=f(g)$ for almost all $(g,h)\in G\times G_0$. Hence, for some Borel set N of G of measure zero, we can assert that for each $g\in G-N$, $f(gh)=f(g)$ for almost all $h\in G_0$. If we write

$$f_1(g) = \int_{G_0} f(gh)d\lambda_0(h),$$

then f_1 is Borel, and for $g\in G-N$, $f_1(g)=f(g)$. Moreover, for any fixed $g\in G-N$ and $h'\in G_0$, the functions $h\to f(gh'h)$ and $h\to f(gh)$ are both equal to $f(g)$ for almost all h. Consequently, $f_1(gh')=f_1(g)$ for *all* $(g,h')\in(G-N)\times G_0$. Consider now the set $X'=\beta[G-N]$. X' is a $\lambda^\sim$ measurable subset of X by corollary 5.2, and hence $\beta^{-1}(X')$ is a λ-measurable set which is a union of left G_0-cosets, and, for all $(g,h)\in\beta^{-1}(X')\times G_0$, $f_1(gh)=f_1(g)$. Moreover, $X-X'$ is null and so $G-\beta^{-1}(X')$ is also null. Define now the function f_2 on G by

$$f_2(g) = \begin{cases} f_1(g) & (g\in\beta^{-1}(X')), \\ 0 & (g\notin\beta^{-1}(X')). \end{cases} \tag{34}$$

f_2 is λ-measurable and $f_2(g)=f(g)$ almost everywhere. Further, f_2 is constant on the left G_0-cosets, and therefore there exists, by corollary 5.20, a $\lambda^\sim$-measurable function f_3 on X such that $f_2=f_3{}^0$. Choose a Borel function a and X such that $a(x)=f_3(x)$ for almost all x. Then $f(g)=a(\beta(g))$ for almost all g. This completes the proof.

5. COCYCLES AND COHOMOLOGY

The functional equations (22) are extremely significant in the theory of induced representations associated with G-spaces. We shall devote the present section to a study of these equations in an appropriately general context.

Let X be a standard Borel G-space, G being a fixed lcsc group. We choose an invariant measure class $\mathscr{C}$ on X. Let M be a standard Borel group, fixed throughout this section. We write 1 for the identity of M. Let f be a function on $G \times X$ with values in M. We shall say that f is a (G,X,M)-*cocycle relative to* $\mathscr{C}$ if the following properties are satisfied:

$$(35)\qquad \begin{aligned} &\text{(i)} && f \text{ is a Borel map of } G\times X \text{ into } M,\\ &\text{(ii)} && f(e,x) = 1 \quad \text{for almost all} \quad x \in X,\\ &\text{(iii)} && f(g_1g_2,x) = f(g_1,g_2\cdot x)f(g_2,x) \quad \text{for almost all}\\ &&& \qquad\qquad (g_1,g_2,x) \in G\times G\times X. \end{aligned}$$

If there is no ambiguity concerning the group M, we shall refer to f simply as a (G,X)-*cocycle relative to* $\mathscr{C}$ or, even more briefly, as a *cocycle relative to* $\mathscr{C}$. Note that the null set in (ii) is with respect to $\mathscr{C}$ while the null sets in (iii) are those of the measure class which is a product of the usual measure classes in G with $\mathscr{C}$.

A cocycle f is said to be a *strict cocycle* if it satisfies (35)(i) and further, (36) (ii′) and (iii′), where

$$(36)\qquad \begin{aligned} &\text{(ii}'\text{)} && f(e,x) = 1 \quad \text{for all} \quad x \in X,\\ &\text{(iii}'\text{)} && f(g_1g_2,x) = f(g_1,g_2\cdot x)f(g_2,x) \quad \text{for all} \quad (g_1,g_2,x) \in G\times G\times X. \end{aligned}$$

Our decision to call the functions satisfying (35) and (36) cocycles is prompted by the fact that these equations are generalizations of the identities which describe the cocycles in the cohomology theory of groups (cf. Eilenberg [1], pp. 3–37). We note here that it was G. W. Mackey who first studied the cocycles in the context of an arbitrary transitive Borel G-space. It was the detailed study and analysis of cocycles which enabled Mackey to formulate and prove the correct generalizations of the classical work of Frobenius on induced representations (of finite groups), to the case when the group in question was an arbitrary lcsc group. The aim of this section is to carry out a cohomology theory of (G,X,M)-cocycles.

We begin by introducing equivalence relations in the set of all cocycles relative to a fixed invariant measure class $\mathscr{C}$. Suppose that f_1 and f_2 are two cocycles relative to $\mathscr{C}$. We shall say that f_1 and f_2 are *cohomologous*, $f_1 \sim f_2$ in symbols, if there exists a Borel map $b(x \to b(x))$ of X into M such that

$$(37)\qquad f_2(g,x) = b(g\cdot x)f_1(g,x)b(x)^{-1}$$

for almost all $(g,x) \in G \times X$. By taking b to be identically 1, we see that if f_1 and f_2 are equal almost everywhere on $G \times X$, then $f_1 \sim f_2$. It is easy to verify that $\sim$ is a genuine equivalence relation in the set of all cocycles relative to $\mathscr{C}$. Each $\sim$ equivalence class will be called a *cohomology class relative to* $\mathscr{C}$. Suppose further that f_1 and f_2 are strict cocycles. We shall then say that f_1 and f_2 are *strictly cohomologous*, $f_1 \approx f_2$ in symbols, if there exists a Borel map $b(x \to b(x))$ of X into M such that

$$f_2(g,x) = b(g \cdot x) f_1(g,x) b(x)^{-1}$$

for *all* $(g,x) \in G \times X$. $\approx$ is also easily verified to be an equivalence relation in the set of all strict cocycles. Each $\approx$-equivalence class will be called a *strict cohomology class*.

The map $\mathbf{1} : g,x \to 1$ is a strict cocycle. Any cocycle cohomologous to it will be called a *coboundary*. If the cocycle is strictly cohomologous to 1, it will be called a *strict coboundary*. Obviously, a cocycle f is a coboundary (strict resp.) if and only if for some Borel function b on X with values in M, $f(g,x) = b(g \cdot x) b(x)^{-1}$ for almost all (all resp.) $(g,x) \in G \times X$.

Suppose that M is abelian. Then the set of Borel maps of $G \times X$ into M is an abelian group under pointwise multiplication; both the set of cocycles and the set of strict cocycles are subgroups of this group. The coboundaries form a subgroup of the group of cocycles and for two cocycles f_1, f_2, $f_1 \sim f_2$ if and only if $f_1 f_2{}^{-1}$ is a coboundary. Therefore the cohomology classes are in one-one correspondence with the elements of the quotient group of the group of cocycles mod the group of coboundaries. This quotient group is known as the *(G,X,M)-cohomology group*.

From now on we shall assume that G is transitive on X. We choose a point $x_0 \in X$ arbitrarily and write G_0 for the stability subgroup at x_0. G_0 is closed. β denotes the map $g \to g \cdot x_0$ of G onto X. β is a Borel map. For any map u of X into some space Z, u^0 denotes the map $g \to u(\beta(g))$ of G into Z. On the other hand, if v is a map of G into Z which is constant on the left G_0-cosets, there exists a unique map u of X into Z such that $u^0 = v$; we denote u by $v^\sim$. Thus $(v^\sim)^0 = v$ and $(u^0)^\sim = u$.

We shall also extend this notation to two variables. Let Y, Z be Borel spaces and let u be a map of $Y \times X$ into Z. We then write u^0 for the map

$$y, g \to u(y, \beta(g))$$

of $Y \times G$ into Z. We note that u is Borel if and only if u^0 is Borel.

At various points of our arguments in this section, we shall be concerned with functions on X, G, $G \times G$, $G \times X$, $G \times G \times X$, and so on. Since G and X possess unique invariant measure classes, there are canonically determined measure classes on each of the other spaces. For instance, if λ and λ' are quasi-invariant σ-finite measures on G and X, respectively, the measure $\lambda \times \lambda'$ determines a measure class on $G \times X$ which depends only on the measure classes of λ and λ'. By a (Borel) null subset of $G \times X$ we mean a set on which all the measures of this class vanish. Similar remarks apply to $G \times G \times X$, etc. In view of the uniqueness of the invariant measure classes, we are entitled to omit any reference to them and speak of cocycles and coboundaries. This we shall do.

Suppose now that f is a strict cocycle. Then it follows easily from (ii′) and (iii′) of (36) that $h \to f(h,x)$ is a Borel homomorphism of the stability group G_x at x into M. We shall call it the *homomorphism defined by f at the point $x \in X$*. If m_1 and m_2 are two Borel homomorphisms of G_0 into M, we shall say that m_1 and m_2 are *equivalent* if there exists a $k \in M$ such that, for all $h \in G_0$,

$$m_2(h) = km_1(h)k^{-1}. \tag{38}$$

The cocycles and their cohomology classes are difficult to handle because of the null sets involved. The strict cocycles and their classes are much better behaved in this respect since there are no null sets. Our main concern is to show that in every cohomology class of cocycles there is a strict cocycle and that the latter is determined uniquely up to strict cohomology. This means that the cohomology classes of cocycles are in canonical one-one correspondence with the strict cohomology classes of strict cocycles, which are easier to analyze.

The simplest strict cocycles are the strict (G,G)-cocycles. Our method of examining strict (G,X)-cocycles consists in "lifting" these to strict (G,G)-cocycles and analyzing the latter.

Lemma 5.23. *Let f be a strict (G,X)-cocycle. Then the function f^0 on $G \times G$ defined by*

$$f^0(g,g') = f(g,\beta(g'))$$

is a strict (G,G)-cocycle. If F is any strict (G,G)-cocycle, there exists a unique Borel map b of G into M such that $b(e)=1$ and

$$F(g,g') = b(gg')b(g')^{-1} \tag{39}$$

for all $(g,g') \in G \times G$.

Proof. The first assertion is trivial. For the second, we note that (36) gives

$$F(g,g') = F(gg',e)F(g',e)^{-1}$$

for all $(g,g') \in G \times G$. Equation (39) follows by putting $b(g')=F(g',e)$. If b_1 is another Borel map of G into M such that $b_1(e)=1$ and $F(g,g')=b_1(gg')b_1(g')^{-1}$, then $b(gg')^{-1}b_1(gg')=b(g')^{-1}b_1(g')$ for all $(g,g') \in G \times G$. This shows that for some $k \in M$, $b(g)^{-1}b_1(g)=k$ for all g. Since $b_1(e)=b(e)=1$, $k=1$ and $b=b_1$.

In view of this lemma, with any strict (G,X) cocycle f, we can associate a Borel function on G. But f^0 is not the most general (G,G)-cocycle since it arises from a (G,X)-cocycle. This Borel function must therefore satisfy some special identities. The next lemma describes these identities.

Lemma 5.24. *Let m be a Borel homomorphism of G_0 into M. Then there exists a Borel map b of G into M such that $b(e)=1$ and*

$$b(gh) = b(g)m(h) \tag{40}$$

for all $(g,h) \in G \times G_0$. Corresponding to any such map b, there is a unique strict (G,X)-cocycle f such that

$$f^0(g,g') = b(gg')b(g')^{-1} \tag{41}$$

for all $(g,g') \in G \times G$; f defines m at x_0. Conversely, if f is a strict (G,X)-cocycle and b is the Borel map of G into M such that $b(e)=1$ and b satisfies (41), *then the restriction of b to G_0 coincides with the homomorphism m defined by f at x_0 and b satisfies* (40).

Proof. Let $m(h \rightarrow m(h))$ be a Borel homomorphism of G_0 into M. Let $c(x \rightarrow c(x))$ be a Borel section for G/G_0. We may assume c so chosen that $c(x_0)=e$. For any $g \in G$, let $a(g)=c(\beta(g))^{-1}g$. Since g and $c(\beta(g))$ lie on the same left coset gG_0, $a(g) \in G_0$. We put

$$b(g) = m(a(g)). \tag{42}$$

Since c and β are Borel maps, $g \rightarrow a(g)$ is a Borel map of G into G_0 and hence $g \rightarrow b(g)$ is a Borel map of G into M. Also $b(e)=m(a(e))=1$. Moreover, for $g \in G$ and $h \in G_0$, $a(gh)=a(g)h$. This shows that

$$b(gh) = b(g)m(h)$$

for all $(g,h) \in G \times G_0$. Let

$$F(g,g') = b(gg')b(g')^{-1}.$$

F is obviously a strict (G,G)-cocycle and (40) shows that $F(g,g'h)= F(g,g')$ for all $(g,g',h) \in G \times G \times G_0$. Hence, there exists a unique Borel map f of $G \times X$ into M such that

$$f(g,\beta(g')) = F(g,g')$$

for all $(g, g') \in G \times G$. It is easily checked that f is a strict (G,X)-cocycle, that $f^0=F$, and that it is the only (G,X)-cocycle with this property. Finally, for $h \in G_0$, $f(h,x_0)=F(h,e)=m(h)$.

For the results in the converse direction, let f be a strict (G,X)-cocycle and let b be such that $b(e)=1$ and

$$f^0(g,g') = b(gg')b(g')^{-1} \tag{43}$$

for all $(g,g') \in G \times G$. Then, for all $(g,g',h) \in G \times G \times G_0$,

$$b(gg'h)b(g'h)^{-1} = b(gg')b(g')^{-1},$$

from which it follows easily that for each $h \in G_0$, $g' \to b(g')^{-1}b(g'h)$ is a constant, say $m(h)$. Thus $b(g'h) = b(g')m(h)$ for all $(g,h) \in G \times G_0$. Since $b(e) = 1$, $b(h) = m(h)$, while (43) implies that $b(h) = f(h,x_0)$, for all $h \in G_0$. This proves everything.

Lemma 5.25. *Let f_i be a strict (G,X)-cocycle and let m_i be the homomorphism defined by f_i at x_0 $(i=1, 2)$. Then the following statements are equivalent:*

$$\text{(44)} \qquad \begin{aligned} &\text{(a)} \quad m_1 \text{ and } m_2 \text{ are equivalent,} \\ &\text{(b)} \quad f_1 \approx f_2, \\ &\text{(c)} \quad f_1 \sim f_2. \end{aligned}$$

Proof. (b) ⇒ (c) is trivial. We shall prove that (a) ⇒ (b) and (c) ⇒ (a).

(a) ⇒ (b). Suppose that m_1 and m_2 are equivalent homomorphisms. Now there exists by lemma 5.24, a Boral map b_i of G into M such that (i) m_i is the restriction of b_i to G_0, (ii) $b_i(gh) = b_i(g)m_i(h)$ for all $(g,h) \in G \times G_0$ and (iii) $f_i{}^0(g,g') = b_i(gg')b_i(g')^{-1}$ for all $(g,g') \in G \times G$. Let $k \in M$ be such that

$$m_2(h) = km_1(h)k^{-1} \tag{45}$$

for all $h \in G_0$. From (45) we obtain the equation

$$b_2(gh)kb_1(gh)^{-1} = b_2(g)kb_1(g)^{-1}$$

for all $(g,h) \in G \times G_0$. This shows that the Borel map $g \to b_2(g)kb_1(g)^{-1}$ is constant on the left G_0-cosets. Therefore there exists a Borel map $b(x \to b(x))$ of X into M with the property that

$$b_2(g)kb_1(g)^{-1} = b(\beta(g)) = b^0(g) \tag{46}$$

for all $g \in G$. An easy computation using (46) and the relation of b_i to f_i now gives us the relation

$$f_2{}^0(g,g') = b^0(gg')f_1{}^0(g,g')b^0(g')^{-1},$$

from which we obtain

$$f_2(g,x) = b(g \cdot x)f_1(g,x)b(x)^{-1}$$

for all $(g,x) \in G \times X$.

(c) ⇒ (a). We assume $f_1 \sim f_2$. Therefore there exists a Borel map b of X into M such that

$$f_2(g,x) = b(g \cdot x)f_1(g,x)b(x)^{-1}$$

for almost all $(g,x) \in G \times X$. This implies that

$$f_2{}^0(g,g') = b^0(gg')f_1{}^0(g,g')b^0(g')^{-1} \tag{47}$$

for almost all $(g,g') \in G \times G$. Let b_i be the Borel map of G into M such that (i) m_i is the restriction of b_i to G_0, (ii) $f_i{}^0(g,g') = b_i(gg')b_i(g')^{-1}$ for all

$(g,g') \in G \times G$ (lemma 5.24). We then obtain from (47), after an easy calculation, the equation

$$b_1(gg')^{-1}b^0(gg')^{-1}b_2(gg') = b_1(g')^{-1}b^0(g')^{-1}b_2(g') \tag{48}$$

for almost all $(g,g') \in G \times G$. This means (corollary 5.21 with $X = G$) that the function $g \to b_1(g)^{-1}b^0(g)^{-1}b_2(g)$ is almost everywhere equal to a constant, say $k^{-1} \in M$. Thus

$$b_2(g) = b^0(g)b_1(g)k^{-1} \tag{49}$$

for almost all $g \in G$. We now conclude from the connection between b_i and m_i that for each $h \in G_0$,

$$b^0(g)b_1(g)k^{-1}m_2(h) = b^0(g)b_1(g)m_1(h)k^{-1}$$

for almost all g. Thus

$$m_2(h) = km_1(h)k^{-1}. \tag{50}$$

This proves that (c) ⇒ (a) and completes the proof.

Lemma 5.26. *If f is any (G,X)-cocycle, there exists a strict (G,X)-cocycle f_1 such that*

$$f(g,x) = f_1(g,x)$$

for almost all $(g,x) \in G \times X$; f_1 is determined uniquely up to strict cohomology.

Proof. Let f be a (G,X)-cocycle and let f^0 be defined as usual, i.e.,

$$f^0(g,g') = f(g,\beta(g')).$$

Then

$$f^0(g_1g_2,g_3) = f^0(g_1,g_2g_3)f^0(g_2,g_3)$$

for almost all $(g_1,g_2,g_3) \in G \times G \times G$. Hence there exists a $g_0 \in G$ such that

$$f^0(g_1g_2,g_0) = f^0(g_1,g_2g_0)f^0(g_2,g_0)$$

for almost all $(g_1,g_2) \in G \times G$. If we put $g = g_1$, $g' = g_2g_0$, we can conclude that

$$f^0(g,g') = f^0(gg'g_0{}^{-1},g_0)f^0(g'g_0{}^{-1},g_0)^{-1}$$

for almost all $(g,g') \in G \times G$. Let

$$d_0(g) = f^0(gg_0{}^{-1},g_0).$$

Then d_0 is a Borel map of G into M and

$$f^0(g,g') = d_0(gg')d_0(g')^{-1}$$

for almost all $(g,g') \in G \times G$. Changing $d_0(g)$ to $d(g) = d_0(g)d_0(e)^{-1}$, we secure $d(e) = 1$ and at the same time

$$f^0(g,g') = d(gg')d(g')^{-1} \tag{51}$$

for almost all $(g,g') \in G \times G$.

Fix now an arbitrary $h \in G_0$. Then it follows from (51) that

$$d(gg'h)d(g'h)^{-1} = d(gg')d(g')^{-1}$$

for almost all $(g,g') \in G \times G$, i.e.,

$$d(gg')^{-1}d(gg'h) = d(g')^{-1}d(g'h)$$

for almost all $(g,g') \in G \times G$. Corollary 5.21 (with $X = G$) implies now that the function $g' \to d(g')^{-1}d(g'h)$ is almost everywhere equal to a constant, say $m(h)$. In other words,

$$d(gh) = d(g)m(h) \tag{52}$$

for almost all g. Clearly $m(h)$ is uniquely determined by the requirement that the condition (52) be satisfied for almost all g. It now follows from (52) and the uniqueness of $m(h)$ that (i) $m(e) = 1$, and (ii) for $h_1, h_2 \in G_0$, $m(h_1h_2) = m(h_1)m(h_2)$. We assert that m is a Borel map of G into M. Since M is standard, we may assume, in this argument, that M is the unit interval. Let $\varphi(g,h) = d(g)^{-1}d(gh)$. Then φ is Borel on $G \times G_0$. Choose a quasi-invariant measure λ on G such that $\lambda(G) = 1$. Then

$$m(h) = \int_G \varphi(g,h)d\lambda(g).$$

As φ is Borel on $G \times G_0$, this shows that m is Borel on G_0. In other words, m is a Borel homomorphism of G_0 into M.

By lemma 5.24 there exists a Borel map d' of G into M such that $d'(e) = 1$ and $d'(gh) = d'(g)m(h)$ for all $(g,h) \in G \times G_0$. From (52) it now follows that for each $h \in G_0$,

$$d(g)d'(g)^{-1} = d(gh)d'(gh)^{-1}$$

for almost all g. By corollary 5.22 we can find a Borel map k of X into M such that $d(g) = k(\beta(g))d'(g)$ for almost all g. Write $d_1(g) = k(\beta(g))d'(g)$. Then $d_1(gh) = d_1(g)m(h)$ for *all* $(g,h) \in G \times G_0$ and $d_1(g) = d(g)$ for almost all g. By lemma 5.24, there exists a unique strict (G,X)-cocycle f_1 such that

$$f_1{}^0(g,g') = d_1(gg')d_1(g')^{-1}$$

for all $(g,g') \in G \times G$. Since $d_1 = d$ almost everywhere, the last equation implies that $f^0 = f_1{}^0$ almost everywhere on $G \times G$. Thus

$$f_1(g,x) = f(g,x) \tag{53}$$

for almost all $(g,x) \in G \times X$.

If f_2 were another strict cocycle such that (53) is satisfied with f_2 in place of f_1, then $f_1 = f_2$ almost everywhere on $G \times X$. In particular, $f_1 \sim f_2$. From lemma 5.25 we now conclude that $f_1 \approx f_2$. This completes the proof.

Theorem 5.27 (*Mackey* [2], [6]). *Let G be a lcsc group, X a standard transitive Borel G-space, and M a standard Borel group, x_0, an arbitrary point of X, and G_0, the stability group at x_0. If γ is any cohomology class of (G,X,M)-cocycles, the set γ_{st} of all strict (G,X,M)-cocycles in γ is nonempty and is a strict cohomology class. The set $\gamma^{\sim}$ of the homomorphisms of G_0 defined by the elements of γ_{st} at x_0 is an equivalence class of Borel homomorphisms of G_0 into M. $\gamma \to \gamma^{\sim}$ is a one-one correspondence of the set of all cohomology classes onto the set of all equivalence classes of Borel homomorphisms of G_0 into M. If M is abelian, the (G,X,M)-cohomology group is isomorphic to the group of all Borel homomorphisms of G_0 into M. Finally, if X is an affine G-space (i.e., G_0 is trivial), all cocycles are coboundaries.*

Proof. The proofs of all the statements are contained in lemmas 5.23 through 5.27.

Remark. The theorem just proved is the cornerstone of the theory of induced representations. It reduces the problem of construction and classification of cocycles to the problem of constructing the equivalence classes of Borel homomorphisms of G_0 into M. Given any Borel homomorphism of G_0 into M, lemma 5.24 tells us the method of constructing the strict cocycles which define this homomorphism. This construction depends however on the choice of a Borel section for the cosets gG_0. *In general, there are no canonical choices for these sections.* We shall see later on how, in at least some special cases, these difficulties can be overcome.

Remark. In many problems, M is already a *topological group* with a second countable topology. It is easy to indicate a sharpened form of theorem 5.27 which can be formulated in this context. We shall say that a map u of a topological space A into a topological space B is *locally bounded* if for any compact set $K \subseteq A$, $u[K]$ has compact closure in B. A mild sharpening of theorem 5.27 is contained in lemma 5.29. To give this we need a preparatory lemma, which is classical.

Lemma 5.28. *Let G be a lcsc group, and M a topological group satisfying the second axiom of countability. Suppose that M is given its natural Borel structure and that m is a homomorphism of G into M which, as a map G into M, is measurable. Then m is continuous.*

Proof. Since we can replace M by the range of m, we shall assume that m maps G onto M. Let U be an open subset of M containing 1 and W an open set such that $1 \in W$ and $WW^{-1} \subseteq U$. Since the open sets sW ($s \in M$) cover M, there is a countable family $\{s_1, s_2, \cdots, s_n, \cdots\}$ such that

$M = \bigcup_n s_n W$. Choose $g_n \in G$ such that $m(g_n) = s_n$. If $m^{-1}(W) = A$, then it is easily seen that $G = \bigcup_n g_n A$. Let μ be a left Haar measure. By assumption, the sets $g_n A$ are all measurable and hence $\mu(A) > 0$. By a well known result, there exists an open set B such that $e \in B \subseteq AA^{-1}$. Since $WW^{-1} \subseteq U$, it follows that $m[B] \subseteq U$. This proves that m is continuous at e. Since m is a homomorphism, m is continuous everywhere.

Lemma 5.29. *Let the notation be as in theorem* 5.27. *Suppose that there is a second countable topology on M which converts M into a topological group and whose associated Borel structure is the given one. Then, any cohomology class γ contains a strict cocycle which is locally bounded as a map of $G \times X$ into M. Moreover, if there is a continuous section for G/G_0, γ contains a strict cocycle which is a continuous map of $G \times X$ into M.*

Proof. It is enough to prove that given a Borel homomorphism m of G_0 into M, there are strict cocycles defining m at x_0 and having the required properties. Now, m, by lemma 5.28, is continuous. Let $c(x \to c(x))$ be a *regular* Borel section for G/G_0 (theorem 5.11) such that $c(x_0) = e$. For $t \in G$, $h(t) = c(\beta(t))^{-1}t$. Then it follows from the regularity of c that h is a locally bounded map from G into G_0. If $b(g) = m(h(g))$ for all $g \in G$, b is a locally bounded map of G into M. $g, g' \to b(gg')b(g')^{-1}$ defines a strict (G,X)-cocycle f; f is clearly locally bounded and defines m at x_0. If c is a continuous map of X into G, f is even continuous.

Remark. The reader would have noticed that the transitivity of the Borel G-space X has played a crucial role in the main theorems of this section. It is a natural question to ask whether some of these results can be generalized to the case when X is not homogeneous.

Suppose then that X is an arbitrary standard Borel G-space and $\mathscr{C}$ an invariant measure class. The measure class will now have to play an explicit role since it is not unique unless X is transitive. With respect to a given measure class $\mathscr{C}$, the problem of describing the cohomology classes of cocycles is then well posed. When there is an orbit $G \cdot x_0$ such that $X - G \cdot x_0$ is a null set for $\mathscr{C}$, this problem is essentially the problem for the transitive Borel G-space $X_1 = G \cdot x_0$. Such measure classes on X we shall call *transitive*.

However, not every invariant measure class is transitive. Let us call a measure class $\mathscr{C}$ which is invariant, an *ergodic measure class* if any Borel function f, having the property that for each $g \in G$, $f(g \cdot x) = f(x)$ for $\mathscr{C}$-almost all x, is constant almost everywhere. Transitive measure classes are ergodic but not conversely. When X is not transitive under G, there will often exist ergodic invariant measure classes $\mathscr{C}$ which are not transitive. For any such $\mathscr{C}$ the problem of describing all the cohomology classes seems quite difficult.

If $\mathscr{C}$ is transitive and G acts freely on X, i.e., for $\mathscr{C}$-almost all $x \in X$ the stability subgroups are *trivial*, i.e., consist only of $\{e\}$, then there exists obviously an $x_0 \in X$ such that $X - G \cdot x_0$ is $\mathscr{C}$-null and $G_{x_0} = e$. In this case, there are no cocycles other than coboundaries. When $\mathscr{C}$ is assumed to be merely ergodic, this is not true any longer. We shall now construct examples where $\mathscr{C}$ is ergodic, G acts freely on X, and yet there are cocycles which are not coboundaries.

Example 1: G is the discrete additive group of integers, X is the multiplicative group of complex numbers of modulus 1, and $M = X$. μ is the Haar measure on X. We denote by z_0 an element of X such that the powers $z_0{}^n$ $(n = 1, 2, \cdots)$ are dense in X. For $n \in G$ and $z \in X$, $n[z] = z_0{}^n z$. The measure class of μ is ergodic, as may be verified easily.

Suppose that f is a Borel map of X into M. If we define, for $(n,x) \in G \times X$,

$$r_f(n,x) = \begin{cases} f(x)f(z_0x)\cdots f(z_0^{n-1}x) & (n \geq 1), \\ 1 & (n = 0), \\ f(z_0{}^{-1}x)^{-1}\cdots f(z_0{}^n x)^{-1} & (n < 0), \end{cases}$$

then r_f is a strict cocycle, as is easily verified, and

$$r_f(1,x) = f(x).$$

It is obvious that every cocycle is equal almost everywhere on $G \times X$ to some r_f. We now ask when r_f is a coboundary. Clearly, this is so if and only if there exists a Borel map b of X into M such that

$$r_f(n,x) = b(z_0{}^n x)b(x)^{-1}$$

for μ-almost all x. In particular,

$$b(x)f(x) = b(z_0x)$$

for almost all x. Now, not every f can satisfy this equation for suitable b. In fact, we have, for $f(x) \equiv x$,

$$\hat{b}_{k+1} = z_0{}^{-k}\hat{b}_k,$$

where $k \rightarrow \hat{b}_k$ is the Fourier transform of b, i.e.,

$$\hat{b}_k = \int_X b(x)x^k d\mu(x), \qquad (k = 0, \pm 1, \pm 2, \cdots).$$

This implies that $|\hat{b}_k|^2$ is independent of k, and hence that $\hat{b}_k = 0$ for all k since $b \in \mathscr{L}^2(X,\mu)$. This means that $b = 0$, a contradiction.

Example 2: In this example, $G = R$ and M the multiplicative group of complex numbers of modulus 1; it is due to Helson and Lowdenslager [1].

We make frequent use of the Pontrjagin theory for locally compact abelian groups (Pontrjagin [1], Kaplansky [2], pp. 54–56).

We write R for the additive group of reals in the usual topology and R^d for the additive group of the reals in its discrete topology. Let D be a *countable* subgroup of R^d and X the multiplicatively written character group of D. In the (usual) topology of pointwise convergence, X is a compact metrizable group. Since no element of D is of finite order, X is connected. For $\lambda \in D$ and $x \in X$ we write $\langle \lambda, x \rangle$ for the value of x at λ. For any $t \in R$, $\lambda \rightarrow e^{it\lambda}$ $(\lambda \in D)$ is an element of X. We write $\hat{t}$ for this element. $t \rightarrow \hat{t}$ is a homomorphism of R into X. We note that the image of R under this homomorphism is dense in X. Otherwise, by the duality there would exist a $\lambda \neq 0$ in D such that

$$e^{it\lambda} = 1$$

for all $t \in R$, which is impossible. If D contains at least two elements which are rationally independent, $t \rightarrow \hat{t}$ is injective, as may be easily verified.

The injection $t \rightarrow \hat{t}$ gives rise to an action of R on X if we set

$$t \cdot x = \hat{t}x \qquad (t \in R,\ x \in X).$$

It is obvious that $t, x \rightarrow t \cdot x$ is a continuous map of $R \times X$ into X and that X becomes an R-space. The Haar measure μ on X is invariant and ergodic under this action. Our object is to exhibit (R,X,M)-cocycles which are not coboundaries.

Suppose $\lambda \in D$. Then the function $\psi_\lambda : t, x \rightarrow \langle \lambda, x \rangle (1 - e^{it\lambda})$ is continuous on $R \times X$ and $\psi_\lambda(t+u, x) = \psi_\lambda(t, u \cdot x) + \psi_\lambda(u, x)$ for all t, $u \in R$ and $x \in X$. However, ψ_λ is not real valued. Hence we pass on to φ_λ, where

$$\varphi_\lambda(t,x) = \langle \lambda, x \rangle (1 - e^{it\lambda}) + \langle -\lambda, x \rangle (1 - e^{-it\lambda}).$$

Then φ_λ is real and $\exp(i\varphi_\lambda)$ is easily verified to be a strict (R,X,M)-cocycle. This cocycle, and hence any finite product of such, is obviously a coboundary. However, things change quite a bit when we take *infinite* products of the cocycles $\exp(i\varphi_\lambda)$. We shall show that many of these infinite products, when convergent, represent cocycles which are not coboundaries.

Let us now assume that $c_1, c_2, \cdots$ and $\lambda_1, \lambda_2, \cdots$ are sequences of positive numbers such that

(a) $c_j \rightarrow +\infty$,

(b) $\lambda_j \in D$ and the λ_j are rationally independent,

(c) $\sum_{j=1}^{\infty} \lambda_j c_j < \infty$.

Put $c_0=0$, $c_j=c_{-j}$, $\lambda_{-j}=-\lambda_j$. Then the series $\sum_{j=-\infty}^{\infty} c_j\langle\lambda_j,x\rangle(1-e^{it\lambda_j})$ converges uniformly when t varies over a bounded set and x varies over X. It therefore represents a real valued function φ continuous on $R\times X$. As

$$\varphi = \sum_{j=1}^{\infty} c_j\varphi_{\lambda_j},$$

$\psi=\exp(i\varphi)$ is a strict cocycle. We shall prove that ψ is not a coboundary. Suppose ψ is a coboundary. Then there exists a Borel map b of X into M such that for

$$(*) \qquad \psi(t,x) = b(\hat{t}x)b(x)^{-1}$$

for almost all $(t,x)\in R\times X$. Write a_t for the map $x\to\psi(t,x)$, and for $y\in X$, let a_y be the map $x\to b(yx)b(x)^{-1}$. Then $y\to a_y$ is a continuous map of X into the Hilbert space $\mathscr{L}^2(\mu)$. This implies that (*) is valid for each t for almost all x.

Suppose now that $y\in X$ is such that $\langle\lambda_j,y\rangle=1$ for all j; then we claim that a_y is a constant (*of absolute value* 1) *for almost all x*. To see this, let us consider, for fixed $s\in R$ and each $y\in X$, the function c_y, where

$$c_y(x) = \exp\left\{(-i)\sum_{j=-\infty}^{\infty} c_j\langle\lambda_j,x\rangle(1-\langle\lambda_j,y\rangle)(1-e^{is\lambda_j})\right\}.$$

Since s is fixed, the series under the exponential is uniformly convergent as x, y vary over X and hence $x, y\to c_y(x)$ is a continuous function on $X\times X$. Hence $y\to c_y$ is a continuous map of X into $\mathscr{L}^2(\mu)$. But if we put $y=\hat{t}$ for $t\in R$, an easy calculation shows that

$$\begin{aligned} c_{\hat{t}}(x) &= \psi(t,s\cdot x)\psi(t,x)^{-1} \\ &= a_{\hat{t}}(s\cdot x)a_{\hat{t}}(x)^{-1}. \end{aligned}$$

The continuity of the maps $y\to c_y$ and $y\to a_y$ shows that for each $y\in X$, we must have

$$c_y(x) = a_y(s\cdot x)a_y(x)^{-1}$$

for almost all x. If now $y\in X$ and $\langle\lambda_j,y\rangle=1$ for all j, $c_y(x)=1$ and hence $a_y(s\cdot x)=a_y(x)$ for almost all x. As R acts ergodically on X, a_y must be a constant almost everywhere. This constant is obviously of modulus 1.

The conclusion of the preceding paragraph can be sharpened. Suppose $y\in X$ is such that $\langle\lambda_j,y\rangle=1$ for $j\neq k$ and $\langle\lambda_k,y\rangle=-1(k, j\geq 1)$. Then, the result of the previous paragraph, applied to the strict cocycle obtained by dropping the terms with indices $\pm k$, shows that for some complex number z of modulus 1,

$$a_y(x) = z\exp\{4ic_k \operatorname{Re}\langle\lambda_k,x\rangle\}$$

for almost all x.

We are now in a position to give the decisive argument. Since the λ_j are rationally independent, there exists, for each $k=1, 2, \cdots$ a $y_k \in X$ such that

$$\langle \lambda_j, y_k \rangle = \begin{cases} 1 & j \neq k, \qquad (j \geq 1), \\ -1 & j = k. \end{cases}$$

Write $a_k{}'$ for a_{y_k}. Since X is compact, there exists a subsequence $\{y_{k_n}\}$ and a $y_0 \in X$ such that $y_{k_n} \rightarrow y_0$. Then $a_{k_n}{}' \rightarrow a_{y_0}$. As $\langle \lambda_j, y_0 \rangle = 1$ for all j, we know that $a_{y_0}(x) = \tau$ for almost all x, τ being a constant with $|\tau| = 1$. Hence

$$\int_X a_{k_n}{}' d\mu \rightarrow \tau.$$

On the other hand, we claim that, since $c_k \rightarrow +\infty$,

$$\int_X a_{y_k} d\mu \rightarrow 0.$$

In fact, consider the integral

$$J_{c,\lambda} = \int_X \exp\{4ic \operatorname{Re}\langle \lambda, x \rangle\} d\mu(x)$$

for $c > 0$. If $\lambda(x) \neq 1$ for some $x \in X$, then $x \rightarrow \langle \lambda, x \rangle$ maps X onto a compact *nontrivial* subgroup of M, which is connected as X is connected, and so coincides with M. The measure μ then goes over to Haar measure on M and hence

$$\begin{aligned} J_{c,\lambda} &= \int_M \exp\{4ic \operatorname{Re} m\} dm \\ &= \frac{1}{2\pi} \int_0^{2\pi} \exp(4ic \cos \theta) d\theta. \end{aligned}$$

$J_{c,\lambda}$ is thus independent of λ and $\rightarrow 0$ as $c \rightarrow \infty$. This proves, as we asserted, that

$$\int_X a_{y_k} d\mu \rightarrow 0.$$

This contradiction shows that our cocycle ψ is not a coboundary.

We shall end this section with an application to the theory of quasi-invariant σ-finite measures on standard transitive Borel G-spaces. For a direct derivation of the results which are about to be obtained now, the reader may consult Mackey's paper [3]. We write R_+ for the multiplicative group of positive reals.

Let G be a lcsc group, G_0 a closed subgroup, $X = G/G_0$ and β the canonical map of G onto X. Let $\beta(e) = x_0$. Suppose that F is a positive

Borel function on $G \times X$ and α is a quasi-invariant σ-finite measure on X. We write

$$F \sim \alpha \tag{54}$$

if for each $g \in G$, $x \to F(g,x)$ is a version of $d\alpha/d\alpha^{(g^{-1})}$.

Theorem 5.30. *Let* Δ *and* Δ_0 *be the modular functions of* G *and* G_0, *respectively. Write*

$$m(h) = \Delta(h)\Delta_0(h)^{-1} \qquad (h \in G_0). \tag{55}$$

Then, for any quasi-invariant σ*-finite measure* α, *there exists a strict* (G,X,R_+)*-cocycle* r *such that* $r \sim \alpha$. *Any such* r *defines the homomorphism* m *at* x_0. *Conversely, if* r *is a strict cocycle defining* m *at* x_0, $r \sim \alpha$ *for some quasi-invariant* σ*-finite measure* α; r *determines* α *up to a constant multiplying factor. Finally, an invariant* σ*-finite measure on* X *is necessarily a Borel measure, and there exists one such on* X *if and only if* $\Delta(h) = \Delta_0(h)$ *for all* $h \in G_0$.

Proof. We begin our proof with a simple observation. Let $S \subseteq G$ be a subgroup such that $G - S$ is measurable and has measure zero. Then $S = G$. In fact, if $S \neq G$, and $g \in G - S$, $gS \subseteq G - S$ so that gS has measure zero. This implies that S is a null set, a contradiction.

This said, let us consider a quasi-invariant σ-finite measure α on X. Let F be a positive Borel function on $G \times X$ such that for almost all g, $x \to F(g,x)$ is a version of $d\alpha/d\alpha^{(g^{-1})}$ (theorem 5.10). Equation (22) shows that F is a (G,X,R_+)-cocycle. Hence, by lemma 5.26 there exists a strict cocycle r such that $r = F$ almost everywhere on $G \times X$. For each $g \in G$, let f_g be a version of $d\alpha/d\alpha^{(g^{-1})}$ and let

$$S = \{g : r(g,x) = f_g(x) \text{ for almost all } x\}.$$

Since the f_g satisfy (20) and r satisfies (36), it is easily seen that S is a *subgroup* of G. Further, $G - S$ has measure zero. Therefore $S = G$. In other words, $r \sim \alpha$.

If r' is another strict cocycle such that $r' \sim \alpha$, then $r' = r$ almost everywhere on $G \times X$. Hence $r' \approx r$ by lemma 5.25. Thus, R_+ being abelian, r' defines the same homomorphism of G_0 as r does, at x_0. Let m_α denote this homomorphism. We claim that m_α is independent of α. Suppose that α' is another σ-finite quasi-invariant measure. Let a be a positive Borel function which is a version of $d\alpha'/d\alpha$. Then it is quickly verified that (with $r \sim \alpha$)

$$r' : g, x \to a(x)r(g,x)a(g \cdot x)^{-1} \tag{56}$$

is a strict cocycle such that $r' \sim \alpha'$. But then, $r' \approx r$, so that $m_{\alpha'} = m_\alpha$. Let m_0 denote this homomorphism of G_0, determined by all the measures α.

Suppose now that r is a strict cocycle defining the homomorphism m_0 at x_0. Let α_1 be a σ-finite quasi-invariant measure and r_1 a strict cocycle such that $r_1 \sim \alpha_1$. Since r_1 also defines m_0 at x_0, $r_1 \approx r$. Hence there exists a Borel map b of X into R_+ such that

$$r(g,x) = b(x)r_1(g,x)b(g\cdot x)^{-1} \tag{57}$$

for all $(g,x) \in G \times X$. If α is the σ-finite quasi-invariant measure such that $d\alpha/d\alpha_1 = b$, then (57) shows that $r \sim \alpha$. If α' is another σ-finite quasi-invariant measure such that $r \sim \alpha'$ also, then α' is a constant multiple of α. In fact, let a be a positive Borel function such that a is a version of $d\alpha'/d\alpha$. Then for each g, $a(x) = a(g\cdot x)$ for almost all x, showing that a is a constant almost everywhere.

Suppose now that X admits an invariant σ-finite measure, say γ. Then the strict cocycle $\mathbf{1} : g, x \to 1$ is such that $\mathbf{1} \sim \gamma$. Since $\mathbf{1}$ defines the trivial homomorphism at x_0, m_0 is trivial. Conversely, let m_0 be trivial and let α be a σ-finite quasi-invariant measure. Let r be a strict cocycle defining m_0 at x_0 such that $r \sim \alpha$. Since m_0 is trivial, $r \approx \mathbf{1}$, and so there exists a Borel map b of X into R_+ such that

$$r(g,x) = b(g\cdot x)b(x)^{-1} \tag{58}$$

for all $(g,x) \in G \times X$. The equation (58) shows that the σ-finite measure γ defined by the equation

$$d\gamma/d\alpha = b$$

is invariant. In other words, X admits a σ-finite invariant measure if and only if m_0 is trivial.

It remains to prove that $m_0 = m$. To this end we use the results and notations of lemmas 5.14 and 5.15 and theorem 5.19. Using the argument of lemma 5.29, we construct a Borel map d of G into R_+, such that (i) $d(e) = 1$ and d is a locally bounded map of G into R_+, (ii) there exists a strict cocycle r such that $r^0(g,g') = d(gg')d(g')^{-1}$ for all $(g,g') \in G \times G$, and (iii) r defines the homomorphism m of G_0 at x_0. Let γ be the Borel measure on G defined by the linear functional (also written γ)

$$\gamma : f \to \int_G f(g)d(g)^{-1}d\mu(g), \qquad (f \in C_c(G))$$

μ being a left Haar measure on G. We claim that γ vanishes on the kernel of the map $f \to f^{\sim}$. In fact, suppose $f \in C_c(G)$ is such that

$$\int_{G_0} f(gh)d\mu_0(h) = 0$$

for all g in G. Multiplying by $u \cdot d^{-1}$ where $u \in C_c(G)$ and integrating with respect to μ, we get

$$\int_G \left(\int_{G_0} f(gh) d\mu_0(h) \right) u(g) d(g)^{-1} d\mu(g) = 0.$$

We interchange the order of integration and change g to gh while integrating with respect to μ. We then obtain, on using the fact that d defines m at x_0,

$$\int_{G_0} \left(\Delta_0(h)^{-1} \int_G u(gh^{-1}) d(g)^{-1} f(g) d\mu(g) \right) d\mu_0(h) = 0.$$

If we now change the order of integration once again we obtain

$$\int_G u^{\sim}(\beta(g)) d(g)^{-1} f(g) d\mu(g) = 0. \tag{59}$$

Since the map $u \to u^{\sim}$ maps $C_c(G)$ onto $C_c(X)$, we can choose u so that $u^{\sim}$ is identically 1 on the set $\beta[K]$, where K is a compact subset of G outside of which f is zero. Then (59) leads to the equation

$$\int_G f(g) d(g)^{-1} d\mu(g) = 0. \tag{60}$$

In other words, there exists a linear functional $\gamma^{\sim}$ on $C_c(X)$ such that $\gamma(f) = \gamma^{\sim}(f^{\sim})$ for all $f \in C_c(G)$. Moreover, we know from lemma 5.15 that given any $v \in C_c(X)$, which is real and ≥ 0, there is a $u \in C_c(G)$ such that $u^{\sim} = v$ and $u \geq 0$. This shows that $\gamma^{\sim}$ is a *nonnegative* linear functional. Therefore there exists a Borel measure α on X, such that $\gamma^{\sim}$ corresponds to α. Then, obviously $\gamma = \alpha^0$. Since γ is quasi invariant, so is α (corollary 5.17). Fix $g \in G$. We shall compute $d\alpha / d\alpha^{(g^{-1})}$. Let t be a positive Borel version of it on X. Then, by corollary 5.17, t^0 is a version of $d\gamma / d\gamma^{(g^{-1})}$. But, since $d\gamma / d\mu = d^{-1}$, we obtain

$$t^0(g') = d(gg') d(g')^{-1} = r^0(g, g')$$

for almost all g'. In other words, for each $g \in G$, $x \to r(g, x)$ is a version of $d\alpha / d\alpha^{(g^{-1})}$. This proves that $r \sim \alpha$. By our definition of m_0, $m = m_0$.

The above argument shows actually that if r is a locally bounded strict cocycle defining m, there exists a quasi-invariant *Borel* measure α such that $r \sim \alpha$. Suppose now that X admits an invariant σ-finite measure α. m is then trivial and our remark at the beginning of this paragraph shows that there exists a *Borel* measure γ such that $\mathbf{1} \sim \gamma$. But this implies that α is a constant multiple of γ by corollary 5.21. α is thus a Borel measure. The proof of the theorem is complete.

Corollary 5.31. *If G is unimodular, G/G_0 admits an invariant Borel measure if and only if G_0 is also unimodular. If G is arbitrary but G_0 is compact, then G/G_0 admits an invariant Borel measure.*

We end this section with an example of a unimodular G and a non-unimodular G_0. Let $G = SL(2,\mathbf{C})$, the group of 2×2 matrices g

$$g = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \qquad (ad - bc = 1) \tag{61}$$

of determinant 1. $g \to a, b, c, d$ gives an injective homeomorphism of G into $\mathbf{C}^4$. The group G, equipped with the relative topology obtained by imbedding it as above in $\mathbf{C}^4$, is a lcsc group. Let G_0 be the subgroup of all elements $t_{a,z}$, where

$$t_{a,z} = \begin{pmatrix} a & z \\ 0 & a^{-1} \end{pmatrix}. \qquad (a, z \in \mathbf{C}, a \neq 0). \tag{62}$$

G_0 is closed. We claim that G is unimodular and G_0 is not, so that $X = G/G_0$ does not admit any σ-finite invariant measure.

For the unimodularity of G we shall be more general and prove that G admits no nontrivial continuous homomorphism into R_+. Suppose that m is one such. Since $m(x) = m(yxy^{-1})$ for all x, y in G, it follows that for any matrix x in G with distinct eigenvalues, $m(x) = m(y)$ for all matrices y with the same eigenvalues as x. Let G' be the set of all matrices x in G with distinct eigenvalues and H the subgroup of diagonal matrices. Then, the restriction of m to H is a continuous homomorphism of H into R_+ with the property that $m(x) = m(y)$ whenever the diagonal entries of x and y are permutations of one another. It is easy to deduce from this that m is identically 1 on H. Since for any $g \in G'$ there is an $a \in G$ such that $aga^{-1} \in H$, $m(g) = 1$ for all $g \in G'$. Now G' is a dense subset of G and so m is identically 1 on G. In particular, G is unimodular.

G_0 is not unimodular. To see this, let

$$H = \{t_{a,0} : a \in \mathbf{C}^*\}. \tag{63}$$

Then H is an abelian closed subgroup of G_0 and is hence unimodular. If G_0 were unimodular, then there would exist an invariant Borel measure on $X_0 = G_0/H$. Now in each left H-coset in G_0 there exists exactly one element of the form $t_{1,z}$ ($z \in \mathbf{C}$); if we use the complex number z to parametrize the coset $t_{1,z}H$, the action of G_0 on X_0 becomes

$$t_{c,z} : \zeta \to c^2\zeta + cz. \tag{64}$$

It is obvious that the space of complex numbers does not admit any Borel measure invariant under the maps (64). Thus G_0 is not unimodular.

It is of interest to note in this case that $X = G/G_0$ is actually *compact*. In fact, if K denotes the subgroup of unitary matrices in G, it is an easy consequence of the existence of the Jordan canonical form of an arbitrary matrix in G, that $G = KG_0$. Hence X is the image of K under the canonical mapping of G. In particular, X is compact.

6. BOREL GROUPS AND THE WEIL TOPOLOGY

Our main concern in this section is to examine how the topology of a lcsc group can be recovered from the measure-theoretic structure. Suppose that G is a lcsc group and that μ is a left Haar measure. Then it is classical that for any Borel set E with $\mu(E)>0$, the set EE^{-1} contains a neighborhood of e. This result hints that the topology of G is completely determined by μ. That this is indeed so was proved by Weil [1] (cf. Halmos [1], pp. 257–289 for a detailed discussion of Weil's results). For our purposes, this result of Weil is somewhat inadequate. The main theorem of this section is a sharpening of Weil's theorem, first proved by Mackey [5]. We shall use this sharpened form of Weil's result in a significant way in Chapter VII.

This sharpening essentially comes out of the following observations. The first point is that starting with only a left invariant measure class on a separable Borel group G one can show that the group possesses a left invariant measure. The second point is that the Weil topology for G is second countable and generates the original Borel structure on G. The third observation now consists in imbedding G as a dense subgroup of a lcsc group G^*. The fact that G is standard now implies that G is a *Borel set* in G^*. This leads easily to the identity $G=G^*$. In other words, every *standard* Borel group admitting a left invariant measure class is actually a lcsc group under its Weil topology. This is the sharpening of Weil's theorem we referred to. It is due to Mackey [5].

Since the main aim of our analysis is to obtain topological results from measure-theoretic assumptions, it is convenient to start with groups without any topology. Throughout this section, G will denote a separable Borel group, i.e., a group G with a Borel structure which converts G into a separable Borel space such that the map $x, y \rightarrow xy^{-1}$ of $G\times G$ into G is Borel. Since the left translations $x \rightarrow zx$ are Borel automorphisms, we may consider measure classes invariant under these. We call them *left invariant*. In an analogous fashion we define *right* invariant measure classes. A measure class which is both left and right invariant is called *bi-invariant*.

Lemma 5.32. *Let G be a separable Borel group on which there exists a left (or right) invariant measure class. Then G admits a bi-invariant measure class.*

Proof. Let $\mathscr{C}$ be a left invariant measure class and λ a finite measure in $\mathscr{C}$. We shall prove first that for any Borel set $E\subseteq G$, $s \rightarrow \lambda(Es)$ is a Borel function of s. In fact, the map $T : s, t \rightarrow s, ts$ is a Borel automorphism of the Borel space $G\times G$ and hence $T[G\times E]$ is a Borel set in $G\times G$. It is then obvious that

$$Es = \{t : (s,t) \in T[G\times E]\},$$

and hence the function $s \rightarrow \lambda(Es)$ is Borel by Fubini's theorem.

Define the measure λ_1 by

$$\lambda_1(E) = \int_G \lambda(Es)d\lambda(s). \tag{65}$$

λ_1 is easily seen to be a measure and $\lambda_1(G) = \lambda(G)^2 > 0$. We claim that the measure class determined by λ_1 is bi-invariant. It is enough to prove that if $\lambda_1(E) = 0$, then $\lambda_1(Ez) = \lambda_1(zE) = 0$ for all $z \in G$. Suppose $\lambda_1(E) = 0$. Then, as $\lambda(Es) \geq 0$, $\lambda(Es) = 0$ for λ-almost all s. Since $\lambda \in \mathscr{C}$ and $\mathscr{C}$ is left invariant $\lambda(zEs) = 0$ for λ-almost all s, showing that $\lambda_1(zE) = 0$. On the other hand, if N is a Borel set such that $\lambda(N) = 0$ and $\lambda(Es) = 0$ for $s \in G - N$, then $\lambda(z^{-1}N) = 0$ and $\lambda(Ezs) = 0$ for $s \in G - z^{-1}N$. This proves that $\lambda_1(Ez) = 0$ also. A similar argument can be given starting with a right invariant measure class. The lemma is proved.

Lemma 5.33. *Let G be a separable Borel group and let $\mathscr{C}$ be a bi-invariant measure class. Then $\mathscr{C}$ contains left invariant and right invariant σ-finite measures.*

Proof. Let λ be a finite measure in $\mathscr{C}$, such that $\lambda(G) = 1$. Define λ^s by

$$\lambda^s(E) = \lambda(s^{-1}E).$$

Then the measures $\{\lambda^s\}$ are all mutually absolutely continuous. Since $s \rightarrow \lambda^s(E)$ is a Borel function on G, corollary 5.6 is applicable and therefore there exists a Borel function r on $G \times G$ such that for λ-almost all s, $t \rightarrow r(s,t)$ is a version of $d\lambda/d\lambda^{(t^{-1})}$. Since λ^t and λ are mutually absolutely continuous, $r(s,t) > 0$ for almost all t for λ-almost all s. Hence r is positive $(\lambda \times \lambda)$-almost everywhere on $G \times G$. We may thus assume that r is positive on $G \times G$.

The arguments of theorem 5.10 can now be used to conclude that for $(\lambda \times \lambda \times \lambda)$-almost all $(s,t,u) \in G \times G \times G$,

$$r(st,u) = r(s,tu)r(t,u).$$

Then there exists a positive Borel function b on G such that

$$r(s,t) = b(st)/b(t) \tag{66}$$

for $(\lambda \times \lambda)$-almost all $(s,t) \in G \times G$. In fact, all we have to do is to imitate the argument of the first part of the proof of lemma 5.26; the only fact used there is the bi-invariance of the class of null sets involved which is certainly satisfied in our present setup. The measure μ such that $d\mu/d\lambda = b$ is then σ-finite and belongs to $\mathscr{C}$. From (66) we conclude that for λ-almost all s, $d\mu/d\mu^{(s^{-1})}$ is equal to 1 almost everywhere. Thus $\mu = \mu^{(s^{-1})}$ for λ-almost all s. Now the set S of all $s \in G$ such that $\mu = \mu^{(s^{-1})}$ is a subgroup. Since this subgroup has a complement of λ-measure zero, it must coincide with G (cf. remark in proof of theorem 5.30). This proves that μ is left invariant. A similar argument yields a right invariant measure in the measure class $\mathscr{C}$. The proof is complete.

Let now $\mathscr{H}$ be a separable Hilbert space and $\mathscr{U}$ the group of all unitary operators on $\mathscr{H}$. We shall give the *strong topology* for $\mathscr{U}$; i.e., the smallest topology which makes all the maps $U \to Uf$ $(f \in \mathscr{H})$ continuous. If $\{U_n\}$ is a sequence of elements in $\mathscr{U}$, $U_n \to V \in \mathscr{U}$ if and only if $U_n f \to Vf$ for all f belonging to some set D such that the linear combinations of vectors of D are dense in $\mathscr{H}$.

Lemma 5.34. *$\mathscr{U}$ is a metrizable topological group satisfying the second axiom of countability and its Borel structure is standard.*

Proof. Since $\mathscr{H}$ is separable, there exists a countable set D dense in $\mathscr{H}$, and the topology for $\mathscr{U}$ is clearly the smallest one which makes all the maps $U \to Uf$ $(f \in D)$ continuous. As there are only denumerably many maps, it follows from standard topological arguments that the topology for $\mathscr{U}$ is metrizable and second countable. To prove that $\mathscr{U}$ is a topological group we must show that the map $U, V \to UV^{-1}$ is continuous. But, as the elements of $\mathscr{U}$ are unitary, we have, for any $f \in \mathscr{H}$,

$$\|UV^{-1}f - U_0V_0{}^{-1}f\| \leq \|V(V_0^{-1}f) - V_0(V_0{}^{-1}f)\| + \|U(V_0{}^{-1}f) - U_0(V_0{}^{-1}f)\|.$$

This inequality proves at once that $\mathscr{U}$ is a topological group.

To prove that the associated Borel structure of $\mathscr{U}$ is standard, we must exhibit a Borel isomorphism of $\mathscr{U}$ with a Borel space which is known to be standard. Let X be the topological product of countably many copies of $\mathscr{H}$ and let $D = \{f_1, f_2, \cdots\}$ be a countable orthonormal basis for $\mathscr{H}$. For each $U \in \mathscr{U}$, let $\zeta(U)$ be the element $(Uf_1, Uf_2, \cdots)$ of X;

$$\zeta : U \to (Uf_1, Uf_2, \cdots). \tag{67}$$

ζ is clearly a homeomorphism of $\mathscr{U}$ into X. We want to show that the range of ζ is a Borel set in X. Clearly, $(z_1, z_2, \cdots) \in X$ lies in the range of ζ if and only if $\{z_1, z_2, \cdots\}$ is an orthonormal basis, i.e., if and only if

$$\begin{aligned} \langle z_m, z_n \rangle &= \delta_{mn} \quad \text{(Kronecker delta)} \\ \sum_m |\langle z_m, f_i \rangle|^2 &= 1 \qquad \text{for all } i. \end{aligned} \tag{68}$$

These are countably many equations and each one of them defines a Borel subset of X. Thus the range of ζ is Borel. ζ is thus an isomorphism of $\mathscr{U}$ with the Borel subspace of X defined by (68). As this is a Borel subset of the standard space X, we can conclude that $\mathscr{U}$ is standard.

Let now G be a separable Borel group and let μ be a left invariant σ-finite measure on G. Let $\mathscr{H}_\mu$ be the Hilbert space of complex valued (equivalence classes of) Borel functions f on G such that

$$\|f\|^2 = \int_G |f(g)|^2 d\mu(g) < \infty. \tag{69}$$

For any $g \in G$, let L_g be the operator of $\mathscr{H}_\mu$ defined by

$$(L_g f)(t) = f(g^{-1}t) \qquad (t \in G). \tag{70}$$

L_g is clearly a unitary operator of $\mathscr{H}_\mu$. We have

$$\begin{aligned} L_e &= 1, && \text{the identity operator,} \\ L_{g_1 g_2} &= L_{g_1} L_{g_2} && \text{for all } g_1, g_2 \in G. \end{aligned} \tag{71}$$

Denote by $\mathscr{U}_\mu$ the group of all unitary operators on $\mathscr{H}_\mu$.

Lemma 5.35. *Let G be a separable Borel group which has a left invariant σ-finite measure μ. Then the map $L(g \to L_g)$ of G into $\mathscr{U}_\mu$ is a one-one Borel map. If G is standard, its range $L[G]$ is a Borel set in $\mathscr{U}_\mu$ and L is a Borel isomorphism of G with $L[G]$.*

Proof. In view of the definition of the Borel structure on $\mathscr{U}_\mu$, we must, in order to show that L is Borel, prove that for each $f \in \mathscr{H}_\mu$, $g \to L_g f$ is a Borel map of G into $\mathscr{H}_\mu$. Since $\mathscr{H}_\mu$ is separable, we can use lemma 5.4 to reduce this to showing that for each $f, f' \in \mathscr{H}_\mu$, $g \to \langle L_g f, f' \rangle$ is a Borel function on G. But ($*$ denotes complex conjugation)

$$\langle L_g f, f' \rangle = \int_G f(g^{-1}t) f'(t)^* d\mu(t),$$

which is obviously Borel, as $g, t \to f(g^{-1}t)f'(t)^*$ is a Borel function on $G \times G$.

We show next that L is one-one. As L is a homomorphism of G into $\mathscr{U}_\mu$, this reduces to showing that if $g \neq e$, $L_g f \neq f$ for some $f \in \mathscr{H}_\mu$. Suppose that $L_g f = f$ for all $f \in \mathscr{H}_\mu$. Since the Borel structure of G is countably generated and μ is σ-finite, there exists a sequence $\{E_n\}$ of sets of finite μ-measure such that $\{E_n\}$ generates the Borel structure of G. Let $f_n = \chi_{E_n}$. Then $f_n \in \mathscr{H}_\mu$ and the sequence $\{f_n\}$ *separates* the points of G (since each single point set of G is a Borel set). Let N be a set of μ-measure zero such that for $x \notin N$,

$$f_n(g^{-1}x) = f_n(x) \tag{72}$$

for all n. The fact that $\{f_n\}$ is a separating sequence now implies that $g^{-1}x = x$ for all $x \in G - N$. The truth of this for even one $x \in G - N$ implies $g = e$. L is thus one-one. The remaining statements follow from the fact that $\mathscr{U}_\mu$ is standard.

We shall now introduce the Weil topology. Let G be a separable Borel group admitting a left invariant σ-finite measure μ. Consider for each $g \in G$, the unitary operator L_g in $\mathscr{H}_\mu$. The *Weil topology for G* is then defined to be the smallest one which makes the maps

$$g \rightarrow L_g f \tag{73}$$

of G into $\mathscr{H}_\mu$ continuous for all $f \in \mathscr{H}_\mu$. In view of the definition of the strong topology for $\mathscr{U}_\mu$, it is clear that the Weil topology is that topology for G which makes the map $g \rightarrow L_g$ of G into $\mathscr{U}_\mu$ a homeomorphism. Thus G, under the Weil topology, is a topological group satisfying the second axiom of countability.

Let E and F be Borel sets of finite measure. We write

$$\rho(E,F) = \mu((E-F) \cup (F-E)). \tag{74}$$

Then from the relations

$$\rho(E,F) = \int_G |\chi_E(t) - \chi_F(t)|^2 d\mu(t),$$

$$\rho(E,F) = \rho(xE,xF),$$

it is easy to deduce that $x_n \rightarrow x$ in the Weil topology for G if and only if $\rho(x_n E, xE) \rightarrow 0$ for all Borel sets E of finite measure; it is even sufficient to restrict the E's to a Boolean ring of sets of finite measure which generate the Borel structure of G.

Lemma 5.36. *If G is a standard Borel group and μ a left invariant σ-finite measure on it, the Borel structure generated by the sets which are open in the Weil topology coincides with the original Borel structure. Moreover, if G is a lcsc group, the Weil topology coincides with the original topology.*

Proof. Let G_1 be the range of $L(g \rightarrow L_g)$. Let $\mathscr{B}^*$ be the Borel structure for G generated by the Weil topology and let $\mathscr{B}$ be the original Borel structure. Then, L being a homeomorphism of G onto G_1, L is a Borel isomorphism of $(G,\mathscr{B}^*)$ with the Borel subspace of $\mathscr{U}_\mu$ defined by G_1. By lemma 5.35, L is also a Borel isomorphism of $(G,\mathscr{B})$, with the Borel subspace of $\mathscr{U}_\mu$ defined by G_1. Hence $\mathscr{B} = \mathscr{B}^*$.

Let us now assume that G is a lcsc group. We want to prove that its topology coincides with the Weil topology. Since both topologies are metrizable, it is obviously sufficient to prove that for any sequence $\{g_n\}$ in G, $g_n \rightarrow e$ in the given topology of G if and only if $\rho(g_n E, E) \rightarrow 0$ for each Borel set E of finite measure. All topological statements, in the argument to follow, refer to the given topology on G. Let then $g_n \rightarrow e$. If $\{V_s\}$ is a decreasing sequence of compact symmetric $(V_s = V_s^{-1})$ neighborhoods of e which form a base for G at e, then for any compact set K, it is easy to see that $\bigcap_s (V_s K) = K$. Fix a Borel set E of finite measure, and let $\varepsilon > 0$. Choose a compact $K \subseteq E$ so that $\mu(E-K) < \varepsilon$ and choose s so that $\mu(V_s K - K) < \varepsilon$. If n_0 is chosen so that $g_n \in V_s$ for $n \geq n_0$, then for all such n,

$$\begin{aligned} \rho(g_n E, E) &\leq \rho(g_n K, K) + 2\rho(E, K) \\ &\leq 2\mu(V_s K - K) + 2\varepsilon \\ &\leq 4\varepsilon. \end{aligned}$$

In the converse direction, suppose $g_n \not\to e$. By passing to a suitable subsequence we may assume that there exists a compact neighborhood V of the identity such that $g_n \notin V$ for all n. Choose a compact neighborhood W of e such that $WW^{-1} \subseteq V$. Then $g_n W \cap W = \varnothing$ for all n so that $\rho(g_n W, W) = 2\mu(W) > 0$. This shows that g_n cannot converge to e in the Weil topology. This completes the proof of the lemma.

Consider now a standard Borel group G with a nonzero left invariant Borel measure μ. We equip G with the Weil topology. Following standard terminology, we shall say that a subset $A \subseteq G$ is *bounded* if for each open set U in G one can find elements $x_1, \cdots, x_n \in G$ such that

$$A \subseteq \bigcup_{i=1}^{n} x_i U;$$

It is obviously enough to require this only for open sets containing e. We shall say that G is *locally bounded* if some open set containing e is bounded. Notice that open sets in G are Borel (lemma 5.36); if U is open and nonempty, there are elements $x_1, x_2, \cdots$, in G such that $G = \bigcup_n x_n U$, from which it follows at once that $0 < \mu(U) \leq \infty$.

Lemma 5.37. *G is locally bounded. Every bounded Borel set is of finite measure.*

Proof. Before we take up the proof of this lemma we make a simple observation. Let E be a Borel set with $E = E^{-1}$ and $0 < \mu(E) \leq \infty$. Then there is a Borel subset $F \subseteq E$ with $F = F^{-1}$ and $0 < \mu(F) < \infty$. In fact, since μ is σ-finite, there is a Borel function f on E with $f(x) > 0$ for all $x \in E$ and $\int_E f d\mu < \infty$; replacing $f(x)$ by $\min(f(x), f(x^{-1}))$ we may assume that $f(x) = f(x^{-1})$ for all x in E; it then suffices to take F as $\{x : f(x) > c\}$ for some $c > 0$.

This said, we come to the proof of lemma 5.37. Let E be a Borel set such that $E = E^{-1}$ and $0 < \mu(E) < \infty$. Put, for $0 < \varepsilon < 2\mu(E)$,

$$N(E : \varepsilon) = \{x : \rho(xE, E) < \varepsilon\}.$$

Since $x \to \rho(xE, E)$ is a continuous function on G, vanishing at e, $N(E : \varepsilon)$ is an open set containing e. We claim that

$$0 < \mu(N(E : \varepsilon)) < \infty.$$

We need only to prove that $\mu(N(E:\varepsilon))<\infty$. Now, we have

$$\iint\limits_{G\times G} \chi_E(t)\chi_E(t^{-1}x)d\mu(x)d\mu(t) = \int\limits_G \chi_E(t)\mu(tE)d\mu(t)$$
$$= \mu(E)^2.$$

On the other hand, as $E=E^{-1}$,

$$\iint\limits_{G\times G} \chi_E(t)\chi_E(t^{-1}x)d\mu(t)d\mu(x) = \iint\limits_{G\times G} \chi_E(t)\chi_E(x^{-1}t)d\mu(t)d\mu(x)$$
$$= \int\limits_G \mu(xE\cap E)d\mu(x).$$

Therefore,

$$\mu(E)^2 \geq \int\limits_{N(E:\varepsilon)} \mu(xE\cap E)d\mu(x).$$

But, if $x\in N(E:\varepsilon)$,

$$\rho(xE,E) = 2\mu(E)-2\mu(xE\cap E) < \varepsilon$$

so that

$$\mu(E)^2 \geq (\mu(E)-\varepsilon/2)\mu(N(E:\varepsilon)).$$

This proves our claim that $0<\mu(N(E:\varepsilon))<\infty$.

At this stage, we know that sufficiently small open subsets of G are of positive finite measure. Let U be any open set containing 1 with $0<\mu(U)<\infty$ and V an open subset of U such that $e\in V$ and $VV^{-1}\subseteq U$. We claim that V is bounded. Otherwise there would exist an open set W containing e and an infinite sequence $x_1, x_2, \cdots$, of elements of V such that

$$x_{n+1}\notin \bigcup_{i=1}^{n} x_iW \qquad (n = 1, 2, \cdots).$$

Choose an open set L containing e such that $LL^{-1}\subseteq W\cap V$. Then it is easy to see that the sets $x_1L, x_2L, \cdots$ are disjoint and contained in VV^{-1}. Since $VV^{-1}\subseteq U$, $0<\mu(U)<\infty$, and since $\mu(L)>0$, we have a contradiction. The last statement is easy. For, let A be a bounded Borel set and W an open set of finite measure. If $x_1, \cdots, x_n\in G$ are such that $A\subseteq\bigcup_{i=1}^n x_iW$, then $\mu(A)\leq n\mu(W)<\infty$.

We shall now use these results to imbed G as a dense subgroup of a lcsc group G^*. Our argument is based on the observation that G admits a left invariant metric metrizing its Weil topology; in fact, if $\{E_n\}$ is a denumerable Boolean ring of Borel sets of finite measure which generate the Borel structure of G, the metric

$$\delta(x,y) = \sum_n \frac{\rho(xE_n,yE_n)}{1+\rho(xE_n,yE_n)}\, 2^{-n} \tag{75}$$

is easily verified to induce the Weil topology; δ is left invariant.

Lemma 5.38. *Let G be a second countable topological group and let it satisfy the two conditions:*

(76)
(a) *there is a left invariant metric δ for the topology of G,*
(b) *if $\{x_n\}$ is a δ-Cauchy sequence, then $\{x_n^{-1}\}$ contains a δ-Cauchy subsequence.*

Let G^* be the completion of the metric space (G,δ) with G considered as imbedded in G^*. Then there exists exactly one way to convert G^* into a topological group such that G becomes a dense topological subgroup of it.

Proof. As $\delta(x_n,x_m)=\delta(x_n^{-1}x_m,e)$, it follows that $\{x_n\}$ is a δ-Cauchy sequence if and only if $x_n^{-1}x_m \to e$ as $n, m \to \infty$. Let G^* be the completion of G (under δ); we regard G as imbedded in G^*. We denote also by δ the extension of the metric δ to G^*, which is a second countable topological space in its metric topology.

We shall now show that if $\{x_n\}$ and $\{y_n\}$ are Cauchy sequences in G, so is $\{x_ny_n\}$. Let V be any open neighborhood of e and let us choose an open neighborhood W of e such that $W^3 \subseteq V$. Let k be an integer such that $y_n^{-1}y_m \in W$ for $n, m \geq k$. Choose an open neighborhood W' of e such that $W' \subseteq W$ and $y_k^{-1}W'y_k \subseteq W$. Finally, select an integer $k_1 \geq k$ such that $x_n^{-1}x_m \in W'$ for $n, m \geq k_1$. Then, if $n, m \geq k_1$,

$$\begin{aligned} y_n^{-1}x_n^{-1} \cdot x_my_m &= (y_n^{-1}y_k) \cdot y_k^{-1}(x_n^{-1}x_m)y_k \cdot y_k^{-1}y_m \\ &\in W(y_k^{-1}W'y_k)W \\ &\subseteq V. \end{aligned}$$

This proves that $\{x_ny_n\}$ is Cauchy. From this it follows easily that there is a unique continuous map p^* of $G^* \times G^*$ into G^* such that $p^*(x,y)=xy$ $(x,y \in G)$. We define, for $x,y \in G^*$,

$$xy = p^*(x,y). \tag{77}$$

This product operation is clearly associative and $xe=ex=x$ for all $x \in G^*$. If $x \in G^*$ and $\{x_n\}$ is any sequence in G converging to x, then, for any limit point x' of the sequence $\{x_n^{-1}\}$ we have $xx'=x'x=e$. Condition (76b) now implies that x^{-1} exists and that x_n^{-1} converges to x^{-1}. The rest of lemma 5.38 now follows trivially.

Corollary 5.39. *Let G be a locally bounded second countable topological group whose topology is induced by a left invariant metric δ. Then G satisfies* (b) *of* (76). *In this case G^* is a lcsc group and for any compact set K of G^*, $K \cap G^*$ is a bounded set in G.*

Proof. Suppose V is a bounded set in G. We claim that any sequence in V has a Cauchy subsequence. To prove this we proceed as follows. Let

$\{U_n\}$ be a sequence of open subsets of G forming a base for the topology of G at e such that $U_{n+1}^{-1}U_{n+1}\subseteq U_n$ for all n. If $\{x_n\}$ is a sequence in V having infinitely many elements, it follows from the boundedness of V that, for each k, some translate a_kU_k contains infinitely many of the x_n. This means that for some infinite set Z_k of the integers, $x_n^{-1}x_m \in U_k^{-1}U_k\subseteq U_{k-1}$ whenever $n, m \in Z_k$. By choosing the Z's to decrease with k and using the Cantor diagonal procedure we may assure ourselves that $\{x_n\}$ has a subsequence which is Cauchy.

We shall now prove that if $\{x_n\}$ is any Cauchy sequence, $\{x_n^{-1}\}$ has a Cauchy subsequence. It is enough to exhibit a bounded set containing all but finitely many of the elements x_n^{-1}. Let U be a bounded neighborhood of e. Now there is a k such that $x_n^{-1}x_k \in U$ for $n \geq k$. Hence $x_n^{-1} \in Ux_k^{-1}$ for all $n \geq k$; as Ux_k^{-1} is bounded, our assertion is proved. G^* can thus be regarded as a second countable topological group and G a dense subgroup. To prove that G^* is locally compact, it is enough to exhibit some compact subset of G^* with nonnull interior. Let V be any subset of G, bounded and open in G and containing e. Since any sequence in V contains a Cauchy subsequence, $\overline{V}$, the closure of V in G^*, is compact. Since G is second countable, there is a sequence $\{x_n\}$ in G such that $G=\bigcup_n x_nV$. We claim that $G^*=\bigcup_n x_n(\overline{V}\,\overline{V})$. In fact, if $y^* \in G^*$ and $\{y_n\}$ is a sequence in G such that $y_n \to y^*$, there exists an m such that $y_m^{-1}y_n \in V$ for all $n \geq m$. Thus $y_n \in y_mV$ for all $n \geq m$. Choose an integer k such that $y_m \in x_kV$. Then $y_n \in x_k\overline{V}\,\overline{V}$ for all $n \geq m$, showing that $y^* \in x_k\overline{V}\,\overline{V}$. By the category theorem of Baire, some $x_k\overline{V}\,\overline{V}$ contains an interior point. This implies that $\overline{V}\,\overline{V}$ is a compact set with nonempty interior and shows that G^* is locally compact.

For the last assertion, let K be compact in G^* and U an open set in G containing e. Let U^* be an open subset of G^* such that $U=U^* \cap G$. Select an open set V^* in G^* such that $e \in V^*$ and $V^*V^*\subseteq U^*$. There exist $x_1^*, x_2^*, \cdots, x_k^* \in G^*$ such that $K\subseteq\bigcup_{i=1}^k x_i^*V^*$. We choose $x_i \in G$ $(1 \leq i \leq k)$ such that $x_i^{-1}x_i^* \in V^*$. Then $K\subseteq\bigcup_{i=1}^k x_iV^*V^*\subseteq\bigcup_{i=1}^k x_iU^*$ and so $K \cap G\subseteq\bigcup_{i=1}^k x_iU$.

We come to our final lemma. Let G be a separable Borel group with a left invariant σ-finite measure μ.

Lemma 5.40. *If G is standard, then G is already locally compact under its Weil topology.*

Proof. Under the Weil topology, G is locally bounded and second countable. We also know that the Weil topology is induced by a left invariant metric. Corollary 5.39 is thus applicable and hence we may regard G as a dense subgroup of a lcsc group G^*. For any Borel set $A^*\subseteq G^*$, $A^* \cap G$ is Borel in G and hence,

$$\mu^* : A^* \to \mu(A^* \cap G) \tag{78}$$

is a measure on G^*. Now μ is finite for bounded sets and so for any compact $K \subseteq G^*$, $\mu^*(K) = \mu(K \cap G) < \infty$. μ^* is thus a Borel measure. We claim that μ^* is left invariant. Clearly μ^* is invariant under left translations by elements of G. Suppose now that $g^* \in G^*$ and $g_n \in G$ such that $g_n \to g^*$. For $f \in C_c(G^*)$ and $h \in G^*$, we write $f^h = L_h f$. Then, for any continuous function f^* on G^* with compact support, $f^{*(g_n)} \to f^{*(g)}$ uniformly while all the functions $f^{*(g_n)}$ vanish outside a fixed compact set. This implies that

$$\int_{G^*} f^{*(g_n)}(t)d\mu^*(t) \to \int_{G^*} f^{*(g)}(t)d\mu^*(t)$$

and hence

$$\int_{G^*} f^*(t)d\mu^*(t) = \int_{G^*} f^{*(g)}(t)d\mu^*(t).$$

Since this is valid for all $f^* \in C_c(G^*)$, it is clear that μ^* is invariant under left translation by g^*. In other words, μ^* is a left Haar measure. μ^* is evidently nonzero.

Now G is standard and the identity map of G into G^* is Borel and one-one. Hence G is a *Borel set* in G^*. But then

$$\begin{aligned}\mu^*(G^* - G) &= \mu((G^* - G) \cap G) \\ &= 0.\end{aligned}$$

Thus G is a subgroup which is also a Borel set with a complement of zero measure in G^*. Therefore, the observation we made at the beginning of theorem 5.30 is applicable and enables us to conclude that $G = G^*$. In other words, G is already locally compact in its Weil topology.

Combining lemmas 5.32 through 5.40 we obtain the following theorem.

Theorem 5.41. (*Mackey* [5]). *Let G be a standard Borel group. If G admits a left (or right) invariant measure class, then G admits a left or right invariant σ-finite measure. Moreover, there exists exactly one topology for G which converts G into a lcsc group and whose corresponding Borel structure is the original one on G.*

NOTES ON CHAPTER V

The measure theoretic study of the orbit spaces of locally compact groups has been a very active subject in recent years because of possible applications to representation theory, number theory, statistical mechanics, and so on. For a survey treating many of these aspects see R. Zimmer, *Ergodic Theory, group representations, and rigidity*, Bull. Amer. Math. Soc., **6** (1982), pp. 383–416.

CHAPTER VI

SYSTEMS OF IMPRIMITIVITY

1. DEFINITIONS

This chapter discusses the notion of a system of imprimitivity for a group and, in some cases, describes all such systems to within unitary equivalence. These problems were first solved by Frobenius for the case of finite groups. For infinite groups their study is more recent. Their first appearance seems to be in a series of examples constructed by Murray and von Neumann [1], [2] in their theory of rings of operators. Anticipating terminology, we can say that the examples of Murray and von Neumann form a class of ergodic systems of imprimitivity. On the other hand, the first work on transitive systems appeared in 1939. In his great 1939 paper Wigner underscored their importance for building up the theory of representations of the inhomogeneous Lorentz group and studied, although in quite an implicit fashion, the systems of imprimitivity associated with this group. Also Gel'fand and Naĭmark [1] based their solution to the problem of describing the unitary representations of the "$ax+b$" group on an analysis of certain natural systems of imprimitivity.

The systematic development of this concept is due to Mackey. In a series of papers (cf. Mackey [7] for all references) he obtained the main *general* results of the subject and pointed out the far-reaching scope of the ideas involved therein. The variety of applications of these results is quite impressive. They enable one, among other things, to obtain the equivalence of matrix and wave mechanics; they lead to a concise description and classification of the relativistic wave equations; they make possible a unified approach to the theory of representations of the commutation rules; also, on the purely mathematical side, they lead to the notion of induced representations of locally compact groups which has played an important role in the theory of group representations; and finally, they have very interesting implications in the theory of invariant subspaces and factorization of analytic functions (cf. Helson [1]).

We begin with the notion of a representation. Let G be a lcsc group and $\mathscr{H}$ a Hilbert space (always separable). A (*unitary*) *representation* of G in $\mathscr{H}$ is a homomorphism

$$U : g \to U_g \tag{1}$$

of G into the group $\mathscr{U}$ of all unitary operators on $\mathscr{H}$ such that for each $f \in \mathscr{H}$, the map

$$g \to U_g f \tag{2}$$

is continuous from G into $\mathscr{H}$. Since $\mathscr{U}$ is second countable in its strong topology (lemma 5.34), lemmas 5.28 and 5.4 show that for the homomorphism $g \to U_g$ to be a representation it is sufficient that

$$g \to \langle U_g f, f' \rangle \qquad (f, f' \in \mathscr{H}) \tag{3}$$

are measurable on G for all $f, f' \in \mathscr{H}$. A representation U in $\mathscr{H}$ is said to be *equivalent* to a representation U' in $\mathscr{H}'$ if there exists a unitary isomorphism

$$W : \mathscr{H} \to \mathscr{H}' \tag{4}$$

of $\mathscr{H}$ onto $\mathscr{H}'$ such that

$$U_g' = W U_g W^{-1} \tag{5}$$

for all $g \in G$. The representation U on $\mathscr{H}$ is said to be *irreducible* if the only closed linear manifolds of $\mathscr{H}$ which are invariant under all the linear operators U_g are 0 and $\mathscr{H}$.

If G is a locally compact *abelian* group, it is well known that its irreducible representations are none other than its characters; i.e., continuous homomorphisms of G into the multiplicative group of complex numbers of modulus 1. Two characters are equivalent if and only if they are identical. The theory of characters was developed in its definitive form by Pontrjagin [1]; this development was used by Weil [1] to obtain the Plancherel formula for locally compact abelian groups.

On the other hand, when G is not necessarily abelian, but *compact*, a definitive theory can again be developed. The beginnings of this theory can be traced at least to the theory of spherical harmonics. But in its modern form, the theory is undoubtedly the creation of Hermann Weyl. Using powerful and original (transcendental) methods he obtained the representations of arbitrary compact groups, his version of the Plancherel formula for such groups (the Peter-Weyl theorem), and rounded off the theory with a succinct description of the representations of compact semisimple Lie groups (cf. his papers [3], [4], [5] in *Mathematische Zeitschrift* and his expositions [1] (to an audience of physicists) and [2]).

The Lorentz group is neither compact nor abelian. The study of its representations must therefore be based on the theory of representations of lcsc groups which are neither compact nor abelian. This theory is vast and quite recent. A survey of some of this recent work is given in Mackey's address [7].

Suppose that G is a lcsc group. Let X be a standard Borel space. We relate G and X by assuming that X is a Borel G-space. Let $\mathscr{H}$ be a separable Hilbert space. A *system of imprimitivity for* G acting in $\mathscr{H}$ is a pair (U,P), where $P(E \to P_E)$ is a projection valued measure on the class of Borel subsets of X (the projections being defined in $\mathscr{H}$), and $U(g \to U_g)$ is a representation of G in $\mathscr{H}$ such that the relations

$$U_g P_E U_g^{-1} = P_{g \cdot E} \tag{6}$$

are satisfied for all $g \in G$ and all Borel sets $E \subseteq X$. The pair (U,P) is said to be *based on* X. The notions of equivalence and irreducibility are defined for systems of imprimitivity in a fashion similar to the corresponding notions for representations. Two systems (U,P) (in $\mathscr{H}$) and (U',P') (in $\mathscr{H}'$) based on the same G-space X are equivalent if there exists a unitary isomorphism W of $\mathscr{H}$ onto $\mathscr{H}'$ such that

$$\begin{aligned} U_g' &= W U_g W^{-1}, \\ P_E' &= W P_E W^{-1} \end{aligned} \tag{7}$$

for all $g \in G$ and all Borel sets $E \subseteq X$. The system (U,P) is said to be *irreducible* if there is no closed linear manifold of $\mathscr{H}$, other than 0 and $\mathscr{H}$, which is invariant under all P_E and all U_g. Note that if U is an irreducible representation, (U,P) is an irreducible system of imprimitivity, but the converse is generally false. If (U^i,P^i) $(i=1, 2, \cdots)$ is a sequence of systems of imprimitivity all based on X, with (U^i,P^i) acting on $\mathscr{H}^i$, then it is natural to define their *direct sum* as the system (U,P) acting in $\mathscr{H}$, where

$$\mathscr{H} = \bigoplus_i \mathscr{H}^i$$

and where

$$\begin{aligned} P_E &: (f_1, f_2, \cdots) \to (P_E^1 f_1, P_E^2 f_2, \cdots), \\ U_g &: (f_1, f_2, \cdots) \to (U_g^1 f_1, U_g^2 f_2, \cdots). \end{aligned}$$

In symbols:

$$(U,P) = \bigoplus_i (U^i,P^i).$$

In order to be able to handle the concepts of irreducibility, equivalence, and so on, it is convenient to transform these geometric concepts into algebraic ones. It is customary to do this with the help of the notion of the commuting ring. We work with representations, but all of our remarks apply also to systems of imprimitivity. Let U be a representation of G in $\mathscr{H}$. The *commuting ring* of U is the set $\mathfrak{A}$ of all bounded operators A in $\mathscr{H}$ such that

$$U_g A = A U_g \tag{8}$$

for all $g \in G$. Obviously $\mathfrak{A}$ is an algebra and, if $A \in \mathfrak{A}$, then $A^* \in \mathfrak{A}$. Moreover it is easily seen that $\mathfrak{A}$ is closed under the weak topology. $\mathfrak{A}$ is thus a von Neumann algebra (cf. Section 4, Chapter VII, Volume I). The structure of $\mathfrak{A}$ determines very completely the geometry of the action of U. For example, a closed linear manifold L is invariant under U if and only if the projection P^L on L belongs to $\mathfrak{A}$. In this case, by restricting each U_g to L we obtain a representation of G in L, called the *subrepresentation of* U defined by L. Thus, every projection in $\mathfrak{A}$ defines a unique subrepresentation of U. This subrepresentation is irreducible if and only if the associated projection, say Q, in $\mathfrak{A}$, is *minimal*, i.e., if $Q' \leq Q$, Q' is a projection and $Q' \in \mathfrak{A}$, then $Q' = Q$ or $Q' = 0$. Thus, U is equivalent to a direct sum of irreducible representations if and only if 1 is a direct sum of minimal projections in $\mathfrak{A}$. U itself is irreducible when and only when the only projections in $\mathfrak{A}$ are 0 and 1. Now if any operator A belongs to $\mathfrak{A}$, all the spectral projections of $(A+A^*)/2$ and $(A-A^*)/2i$ belong to $\mathfrak{A}$. Thus U is irreducible if and only if $\mathfrak{A}$ consists of the multiples of the identity operator.

These remarks, which will suffice for our purposes, are the basic ones around which a theory of the *global geometry* of an arbitrary representation can be built. The reader, who is interested in this should consult the works of Mackey [4], [7], Varadarajan [5], and Loomis [3].

Suppose (U, P) is a system of imprimitivity based on X acting in $\mathscr{H}$. For any bounded Borel function f on X (complex valued) let the operator A_f be defined by

$$\langle A_f u, v\rangle = \int_X f(x)\,d\nu_{u,v}(x), \tag{9}$$

where $\nu_{u,v}$ is, as usual, the complex valued measure on X defined by

$$\nu_{u,v}(E) = \langle P_E u, v\rangle \qquad (u, v \in \mathscr{H}).$$

$\nu_{u,u}$ is a finite measure for any $u \in \mathscr{H}$ and $\nu_{u,u}(X) = \|u\|^2$. Moreover,

$$\|A_f u\|^2 = \int_X |f(x)|^2 d\nu_{u,u}(x) \qquad (u \in \mathscr{H}).$$

The mapping $f \to A_f$ from the algebra of all bounded Borel functions into the algebra of all bounded operators in $\mathscr{H}$ has the properties described below; their verifications are easy:

$$\text{(10)}\quad \begin{array}{rl} \text{(i)} & f \to A_f \text{ is an algebra homomorphism,} \\ \text{(ii)} & \text{if } f \text{ is real, } A_f \text{ is self-adjoint; more generally } A_f{}^* = A_{f^*}, \\ \text{(iii)} & \|A_f\| \leq \sup\limits_{x \in X} |f(x)|. \end{array}$$

Moreover, using standard arguments, we obtain, from (6),

$$U_g A_f U_{g^{-1}} = A_{(f^g)}, \tag{11}$$

where f^g is the function $x \to f(g^{-1} \cdot x)$.

An important special case arises when X is a locally compact Hausdorff space satisfying the second axiom of countability and the action of G on X is continuous. In this case, it is perhaps more natural to replace the projection valued measure P by the homomorphism $f \to A_f$ $(f \in C_c(X))$. We have:

Lemma 6.1. *Let $f \to A_f$ be a homomorphism of the algebra $C_c(X)$ into the algebra of all bounded operators on a separable Hilbert space $\mathscr{H}$ satisfying the relations* (10) *and such that the linear manifolds spanned by the ranges of all the operators A_f is dense in $\mathscr{H}$. Suppose further that $g \to U_g$ is a unitary representation of G in $\mathscr{H}$ such that (U,A) satisfies* (11). *Then, there exists exactly one system of imprimitivity (U,P) based on X such that, for each $f \in C_c(X)$, A_f is related to P by* (9).

Proof. For any f in $C_c(X)$ which is real and ≥ 0, we observe that $A_f \geq 0$; in fact, $f = f_1{}^2$, where f_1 is also real and ≥ 0 and $A_f = A_{f_1}{}^2 \geq 0$. Consequently, the map $f \to \langle A_f u, u\rangle$ defines a Borel measure $\nu_{u,u}$ on X. If K is any compact set and f is an element of $C_c(X)$ such that $0 \leq f \leq 1$ and $f = 1$ on K, the inequality (iii) of (10) shows that

$$\nu_{u,u}(K) \leq \|u\|^2.$$

$\nu_{u,u}$ is thus a finite measure and $\nu_{u,u}(X) \leq \|u\|^2$. At the same time, if $u, v \in \mathscr{H}$ there exists a unique complex valued measure $\nu_{u,v}$ such that

$$\langle A_f u, v\rangle = \int_X f d\nu_{u,v} \qquad (f \in C_c(X)). \tag{12}$$

For any Borel set E, the map $u, v \to \nu_{u,v}(E)$ is $*$-bilinear and $0 \leq \nu_{u,u}(E) \leq \|u\|^2$ for all $u \in \mathscr{H}$. Therefore, there exists a self-adjoint operator P_E with $0 \leq P_E \leq 1$ such that

$$\nu_{u,v}(E) = \langle P_E u, v\rangle \tag{13}$$

for all $u, v \in \mathscr{H}$.

We want to show that $E \to P_E$ is a projection valued measure. Suppose K is a compact set and $\{f_n\}$ is a sequence of real nonnegative functions in $C_c(X)$ such that $f_n \downarrow \chi_K$ pointwise. Then (12) shows that

$$\langle A_{f_n} u, v\rangle \to \langle P_K u, v\rangle \tag{14}$$

for all $u, v \in \mathscr{H}$; i.e., A_{f_n} converges weakly to P_K ($\Rightarrow$ will denote weak convergence). If f is any element of $C_c(X)$, $A_{f_n} A_f \Rightarrow P_K A_f$ and $A_f A_{f_n} \Rightarrow A_f P_K$ so that P_K commutes with all A_f. Moreover, if $f \geq 0$, then $f f_n \downarrow f \chi_K$ and hence

$$\begin{aligned} \langle A_{f_n} A_f u, v\rangle &= \int_X f f_n d\nu_{u,v} \\ &\to \int_K f d\nu_{u,v}. \end{aligned}$$

In other words, if $f \in C_c(X)$ is real and ≥ 0,

$$\langle P_K A_f u, v \rangle = \int_K f d\nu_{u,v}. \tag{15}$$

If we choose another compact set L and a sequence $\{f_n'\}$ of real non-negative elements of $C_c(X)$ such that $f_n' \downarrow \chi_L$, we obtain, from (15) and the convergence $A_{f_n'} \Rightarrow P_L$, the equation

$$\langle P_L u, P_K v \rangle = \langle P_{K \cap L} u, v \rangle.$$

Now, for fixed K, $E \to \langle P_E u, P_K v \rangle$ and $E \to \langle P_{E \cap K} u, v \rangle$ are complex valued measures which coincide for all compact sets L. Hence they coincide for all E. For arbitrary but fixed E, we now use the resulting coincidence for all K, and conclude that for all Borel sets E and F,

$$\langle P_E u, P_F v \rangle = \langle P_{E \cap F} u, v \rangle$$

for all $u, v \in \mathscr{H}$, i.e., $P_{E \cap F} = P_E P_F$. For $E = F$ this shows that P_E is a projection. The map $E \to P_E$ has now all the properties of a projection valued measure except perhaps that P_X may not be 1. But $P_X = 1$; for, if $P_X \neq 1$, there will be a nonzero $u \in \mathscr{H}$ such that $P_X u = 0$. This would imply that $\nu_{u,u} = 0$ and hence that $A_f{}^* u = 0$ for all $f \in C_c(X)$. But then u would be orthogonal to the ranges of all the A_f, a contradiction. Finally, by a routine transport of structure, the equation

$$U_g A_f U_g{}^{-1} = A_{(f^g)}$$

is seen to lead to the equation

$$U_g P_E U_g{}^{-1} = P_{g \cdot E}.$$

This proves the lemma.

We shall now give some examples of systems of imprimitivity. X and G have the usual meanings, i.e., X is a standard Borel G-space, G is lcsc.

Example 1: Let α be a σ-finite invariant measure on X, $\mathscr{H} = \mathscr{L}^2(X, \alpha)$; U_g and P_E are defined by

$$\begin{aligned} P_E f &= \chi_E f, \\ (U_g f)(x) &= f(g^{-1} \cdot x). \end{aligned} \tag{16}$$

The identities $U_{g_1 g_2} = U_{g_1} U_{g_2}$, $U_e = 1$ as well as the relations (6) are the results of routine calculations. Further, the invariance of α implies that U_g is unitary. For $f, f' \in \mathscr{H}$, the map $g, x \to f(g^{-1} \cdot x) f'(x)^*$ is Borel on $G \times X$ and hence the map

$$g \to \langle U_g f, f' \rangle = \int_X f(g^{-1} \cdot x) f'(x)^* d\alpha(x)$$

is a Borel map. U is thus a representation and (U,P) a system of imprimitivity.

Example 2: Our second example is a variation on the theme of the first one. Let $\mathscr{K}$ be a separable Hilbert space. We then define, for any σ-finite measure α on X, $\mathscr{H}=\mathscr{L}^2(X,\mathscr{K},\alpha)$, to be the vector space of α-equivalence classes of Borel maps f of X into $\mathscr{K}$ such that

$$\|f\|^2 = \int_X |f(x)|^2 d\alpha(x) < \infty; \tag{17}$$

here $|\cdot|$ denotes the norm in $\mathscr{K}$. If we define, for $f, f' \in \mathscr{H}$,

$$\langle f : f'\rangle = \int_X \langle f(x),f'(x)\rangle d\alpha(x), \tag{18}$$

where $\langle . , . \rangle$ denotes the inner product in $\mathscr{K}$, then $\mathscr{H}$ becomes a Hilbert space under $\langle . : . \rangle$. If X is a separable Borel space, $\mathscr{H}$ is separable.

We shall now select an arbitrary invariant measure class on X and a σ-finite measure α in it. *α need not be invariant.* For each $g \in G$ we select a positive Borel function r_g which is a version of $d\alpha/d\alpha^{(g^{-1})}$. We then define, for $f \in \mathscr{H}$

$$\begin{aligned} P_E f &= \chi_E f, \\ U_g f &= \{(r_g)^g\}^{1/2} f^g, \end{aligned} \tag{19}$$

i.e.,

$$(U_g f)(x) = \{r_g(g^{-1}\cdot x)\}^{1/2} f(g^{-1}\cdot x). \tag{20}$$

The quasi invariance of α implies that everything in sight is well defined. The identities (20) of Chapter V, which are atisfied by the r_g, now lead to the conclusion that $g \to U_g$ is a homomorphism of G and that (U,P) satisfies (6). The calculations are straightforward. Moreover, as each U_g is invertible and

$$\begin{aligned} \|U_g f\|^2 &= \int_X (r_g)^g |f^g|^2 d\alpha(x) \\ &= \int_X r_g |f(x)|^2 d\alpha^{(g^{-1})}(x) \\ &= \|f\|^2, \end{aligned}$$

it follows that each U_g is a unitary operator. In order to assure ourselves that (U,P) is a system of imprimitivity, it remains only to check that for each $f, f' \in \mathscr{H}$, $g \to \langle U_g f : f'\rangle$ is a measurable function on G. Now, by theorem 5.10, there exists a positive Borel function r on $G \times X$ such that for almost all g, $r_g(x)=r(g,x)$ for α-almost all x. Therefore, for almost all g,

$$\langle U_g f : f'\rangle = \int_X \{r(g,g^{-1}\cdot x)\}^{1/2} \langle f(g^{-1}\cdot x), f'(x)\rangle d\alpha(x).$$

The right side is a Borel function of g since the integrand is Borel on $G \times X$, so that $g \to \langle U_g f : f' \rangle$ is a measurable function. But, as noted above, this is sufficient to assure us that U is a representation.

When X is the phase space of a classical mechanical system, $G = R^1$, and the action of G on X is given by the dynamical group of the system, the representation U in example 1 was considered by Koopman [1] (α is the Liouville measure). Prompted by this, we shall call the system of example 2 (of which the system in example 1 is a special case) a *Koopman system of imprimitivity.*

In this example, the fact that (U,P) satisfies (6) depends decisively on the identities satisfied by the r_g. It is natural to suspect that one may be able to construct other systems of imprimitivity, by using more general cocycles. This is substantially the case. We shall develop this idea in the following sections.

2. HILBERT SPACES OF VECTOR VALUED FUNCTIONS

The Hilbert space in which the system of imprimitivity of example 2 of Section 1 acts is a space of $\mathscr{K}$-valued functions. This situation is very typical. We devote the present section to a detailed study of certain algebras of operators in such Hilbert spaces.

Let $\mathscr{K}$ be a separable Hilbert space. We use the notation $|\cdot|$ to denote the norm in $\mathscr{K}$ and $\langle . , . \rangle$ to denote the inner product in $\mathscr{K}$. For any operator A of $\mathscr{K}$ we write $|A|$ for its norm. Let S be the set of all bounded operators in $\mathscr{K}$. The unitary group M of $\mathscr{K}$ is a subset of S. We shall consider S as a Borel space by giving it the smallest Borel structure which makes all the maps $A \to \langle Au,v \rangle$ $(u, v \in \mathscr{K})$ Borel. When we speak of Borel maps from or into S, this is the Borel structure we have in mind even when it is not made explicit.

Lemma 6.2. *S, equipped with the above Borel structure, becomes a standard Borel space and M is a Borel subset of it. The maps $A, B \to AB$ and $A, B \to A \pm B$ are Borel from $S \times S$ into S, $A, u \to Au$ is Borel from $S \times \mathscr{K}$ into $\mathscr{K}$, while $A \to A^*$ is a Borel automorphism of S. The induced Borel structure on M coincides with the Borel structure associated with the strong topology on M.*

Proof. Choose an orthonormal basis $\{e_n\}$ in $\mathscr{K}$. Let T be the space of all complex matrices (a_{ij}). We may identify T canonically with a product of copies of the complex plane. T thus becomes a standard Borel space. For $A \in S$, we define

$$a_{ij}(A) = \langle Ae_j, e_i \rangle. \tag{21}$$

Obviously, the Borel structure on S is the smallest one which makes all the maps $a_{ij}(A \to a_{ij}(A))$ Borel. Thus, if $t(A)$ denotes the element $(a_{ij}(A))$ of T, $t(A \to t(A))$ is a one-one Borel map of S onto a set, say S' contained in T, and the Borel structure on S is that one for which t is a Borel isomorphism with the Borel subspace of T defined by S'. To prove that S is standard, it is sufficient to prove that S' is a Borel subset of T. Only the case dim $\mathscr{K} = \infty$ is of interest. Now, if (a_{ij}) is an element of T, it is easily seen to belong to S' if and only if there is an integer L such that for any integer N, and any N complex numbers $c_1 \ldots, c_N$ with rational real and imaginary parts,

$$\sum_i \left| \sum_{j=1}^{N} a_{ij} c_j \right|^2 \leq L(|c_1|^2 + \cdots + |c_N|^2).$$

Since the conditions involved are only denumerably many and since each condition defines a Borel subset of T, it follows that S' is a Borel set. This proves that S is standard.

The equations

$$\begin{aligned} a_{ij}(AB) &= \sum_k a_{ik}(A) a_{kj}(B), \\ a_{ij}(A^*) &= a_{ji}(A)^*, \\ a_{ij}(A \pm B) &= a_{ij}(A) \pm a_{ij}(B), \\ \langle Au, e_k \rangle &= \sum_j \langle u, e_j \rangle a_{kj}(A), \end{aligned}$$

tell us immediately that $A, u \to Au$, $A, B \to AB$, and $A \to A^*$ are Borel maps with appropriate domains and ranges. Since

$$M = \{A : A \in S, AA^* = A^*A = 1\},$$

it follows from these results that M is a Borel set in S. Finally, the induced Borel structure on M is the smallest one which makes all the maps $A \to Au$ of M into $\mathscr{K}$ $(u \in \mathscr{K})$ Borel, and is thus clearly the Borel structure associated with the strong topology on M.

Corollary 6.3. *Suppose Z is a Borel space and $z \to A(z)$, $z \to B(z)$, and $z \to u(z)$ are Borel mappings of Z into S, S, and $\mathscr{K}$, respectively. Then $z \to A^*(z)$ and $z \to A(z)B(z)$ are Borel maps of Z into S and $z \to A(z)u(z)$ is a Borel map of Z into $\mathscr{K}$.*

Consider now a standard Borel space X, a σ-finite measure α on X and the Hilbert space $\mathscr{H} = \mathscr{L}^2(X, \mathscr{K}, \alpha)$. For each Borel set $E \subseteq X$,

$$P_E : f \to \chi_E f \qquad (f \in \mathscr{H}) \tag{22}$$

is a projection operator, and $P(E \to P_E)$ is a projection valued measure. The main concern of this section is with the algebra of bounded operators of $\mathscr{H}$ which commute with all P_E. We consider *bounded* Borel maps b:

$$b : x \to b(x), C = \sup_{x \in X} |b(x)| < \infty \tag{23}$$

of X into S. If f is any Borel function on X with values in $\mathscr{K}$ such that $\int_X |f(x)|^2 d\alpha(x) < \infty$, then $x \to b(x)f(x)$ is also Borel and

$$\int_X |b(x)f(x)|^2 d\alpha(x) \leq C \int_X |f(x)|^2 d\alpha(x).$$

Thus there exists a unique bounded operator, denoted by $b^\sim$, such that for each $f \in \mathscr{H}$ for α-almost all x,

$$(b^\sim f)(x) = b(x)f(x). \tag{24}$$

$b^\sim$ is an operator in $\mathscr{H}$, its norm $\|b^\sim\| \leq C$, and it follows from (24) that $b^\sim$ commutes with P, i.e., for all E,

$$b^\sim P_E = P_E b^\sim. \tag{25}$$

It is obvious that if b_1 is another bounded Borel map of X into S such that $b_1(x) = b(x)$ for α-almost all x, $b_1{}^\sim = b^\sim$. We regard the collection of bounded Borel maps from X into S as an *algebra with involution* by defining

$$\begin{aligned} (a_1 b_1 + a_2 b_2)(x) &= a_1 b_1(x) + a_2 b_2(x) \qquad (a_1, a_2 \text{ complex numbers}) \\ (bc)(x) &= b(x)c(x), \\ b^*(x) &= b(x)^*, \end{aligned} \tag{26}$$

for all $x \in X$.

Lemma 6.4. *For any bounded Borel map b of X into S, $b^\sim$ is a bounded operator in $\mathscr{H}$ commuting with P. Conversely, if B is an operator in $\mathscr{H}$ which commutes with P, there exists a bounded Borel map b of X into S such that $b^\sim = B$. $b_1{}^\sim = b_2{}^\sim$ if and only if $b_1(x) = b_2(x)$ for α-almost all x. Finally, $(b_1 b_2)^\sim = b_1{}^\sim b_2{}^\sim$, $(b_1{}^*)^\sim = (b_1{}^\sim)^*$. In particular, $b^\sim$ is a unitary operator (projection, normal, etc.) if and only if $b(x)$ is unitary (projection, normal, etc.) in $\mathscr{K}$ for α-almost all x.*

Proof. We assume that $\alpha(X) < \infty$; the general case is obtained by an obvious patching up over sets of finite measure which we leave to the reader. For any $u \in \mathscr{K}$ we write, for the purpose of this proof, $\mathbf{u}$ for the constant function $x \to u$ of X into $\mathscr{K}$; $\mathbf{u} \in \mathscr{H}$ as $\alpha(X) < \infty$.

We observe that if B_1 and B_2 are two bounded operators of $\mathscr{H}$ which commute with P, and if $B_1 \mathbf{u} = B_2 \mathbf{u}$ for all $u \in \mathscr{K}$, then $B_1 = B_2$. In fact, for any Borel set $E \subseteq X$, $B_1(\chi_E \mathbf{u}) = B_2(\chi_E \mathbf{u})$. From this it follows that $B_1 f = B_2 f$ whenever f is a Borel function on X (with values in $\mathscr{K}$) taking only finitely many values. The general case now follows from the fact that such functions are dense in $\mathscr{H}$. Thus $B_1 = B_2$.

This done, let B be a bounded operator in $\mathscr{H}$ such that B commutes with P. We may assume that $\|B\| \leq 1$. Let $\{e_n\}$ be an orthonormal basis for $\mathscr{K}$. Let b_n be a Borel map of X into $\mathscr{K}$ such that $b_n = B\mathbf{e}_n$. Now $\|B\| \leq 1$; so, for any N complex numbers $c_1, \cdots, c_N$, as B commutes with P,

$$\begin{aligned} \|BP_E(c_1\mathbf{e}_1 + \cdots + c_N\mathbf{e}_N)\|^2 &= \|\chi_E(c_1 B\mathbf{e}_1 + \cdots + c_N B\mathbf{e}_N)\|^2 \\ &\leq \|\chi_E(c_1\mathbf{e}_1 + \cdots + c_N\mathbf{e}_N)\|^2, \end{aligned}$$

i.e.,

$$\|\chi_E(c_1 B\mathbf{e}_1 + \cdots + c_N B\mathbf{e}_N)\|^2 \leq (|c_1|^2 + \cdots + |c_N|^2)\alpha(E)$$

for all E, i.e.,

$$\int_E |c_1 b_1(x) + \cdots + c_N b_N(x)|^2 d\alpha(x) \leq (|c_1|^2 + \cdots + |c_N|^2)\alpha(E)$$

for all E. For fixed N and $c_1, \cdots, c_N$, these inequalities (for all E) imply that

$$|c_1 b_1(x) + \cdots + c_N b_N(x)|^2 \leq |c_1|^2 + \cdots + |c_N|^2$$

for α-almost all x. By varying N and the c's over complex numbers with rational real and imaginary parts, we can assure ourselves that for some α-null set Z we have the inequalities

$$|c_1 b_1(x) + \cdots + c_N b_N(x)|^2 \leq |c_1|^2 + \cdots + |c_N|^2 \tag{27}$$

valid for all $x \in X - Z$, all integers N, and all sets $(c_1, \cdots, c_N)$ of N-tuples of complex numbers with rational real and imaginary parts. From (27) we conclude that for each $x \in X - Z$ there exists an operator $b(x)$ in $\mathscr{K}$ with $|b(x)| \leq 1$ such that

$$b_r(x) = b(x)e_r \tag{28}$$

for all r. We define $b(x) = 0$ for $x \in Z$. Equation (28) shows at once that $x \to b(x)$ is a Borel map from X into S such that $|b(x)| \leq 1$ for all x and that for the operator $b^{\sim}$

$$b^{\sim}\mathbf{e}_r = B\mathbf{e}_r \tag{29}$$

for all r. From (29) and from the observation made at the beginning of the proof, we deduce that $B = b^{\sim}$. To prove that b is unique up to null sets, let b_1 be such that $b_1{}^{\sim} = b^{\sim}$. Then for each r, $b_1(x)e_r = b(x)e_r$ for α-almost all x. This shows at once that $b_1(x) = b(x)$ for α-almost all x.

Routine calculations show that $(bc)^\sim = b^\sim c^\sim$, $(b \pm c)^\sim = b^\sim \pm c^\sim$, and $(b^*)^\sim = b^{\sim *}$. If $b(x)$ is unitary for almost all x, then it is obvious that $b^\sim$ is unitary. Conversely, if $b^\sim b^{\sim *} = b^{\sim *} b^\sim = 1$, then $b(x)b(x)^* = b(x)^* b(x) = 1$ for almost all x, by the essential one-one nature of the map $b \to b^\sim$. Thus $b(x)$ is unitary for almost all x. The assertions involving projection and normal operators are proved similarly. The proof of the lemma is complete.

Our next lemma deals with families of operators commuting with P.

Lemma 6.5. *With notation above, let Y be a standard Borel space and*

$$b : y, x \to b(y,x) \tag{30}$$

a bounded Borel map of $Y \times X$ into S. For each $y \in Y$, let $b(y,\cdot)$ denote the map $x \to b(y,x)$ of X into S. Then the map $y \to b(y,\cdot)^\sim$ is a bounded Borel map of Y into the Borel space of all bounded operators of $\mathscr{H}$. Conversely, if $y \to B_y$ is a bounded Borel map of Y into the Borel space of all bounded operators in $\mathscr{H}$ such that B_y commutes with P for all y, and if ν is a σ-finite measure on Y, there exists a bounded Borel map b of $Y \times X$ into S such that

$$B_y = b(y,\cdot)^\sim \tag{31}$$

for ν-almost all y. b is determined uniquely $(\nu \times \alpha)$-almost everywhere. If each B_y is unitary, we can arrange matters so that $b(y,x)$ is unitary for all y, x.

Proof. We may clearly assume that α and ν are finite. Let b be a bounded Borel map of $Y \times X$ into S, say $|b(y,x)| \leq C$, and let $B_y = b(y,\cdot)^\sim$. If f and f' are in $\mathscr{H}$,

$$\langle B_y f : f' \rangle = \int_X \langle b(y,x) f(x), f'(x) \rangle d\alpha(x),$$

which is Borel in y as the integrand is Borel in y, x. We also have $\|B_y\| \leq C$. Conversely, let $y \to B_y$ be a bounded Borel map of Y into the Borel space of all bounded operators of $\mathscr{H}$ and let B_y commute with P for all y. We may assume that $\|B_y\| \leq 1$ for all y. For each $y \in Y$ we choose a Borel map b_y of X into S such that $|b_y(x)| \leq 1$ for all x and $B_y = b_y^\sim$. Let $\{e_n\}$ be an orthonormal basis for $\mathscr{K}$ and for any $u \in \mathscr{K}$ let us write $\mathbf{u}$ for the constant map $x \to u$ on X. Then, for $E \subseteq X$,

$$\langle B_y \chi_E \mathbf{e}_m : \mathbf{e}_n \rangle = \int_E \langle b_y(x) e_m, e_n \rangle d\alpha(x),$$

so that for each Borel set $E \subseteq X$,

$$y \to \int_E \langle b_y(x) e_m, e_n \rangle d\alpha(x)$$

is a Borel function of y. Lemma 5.5 is now applicable and hence we can deduce the existence of complex-valued Borel functions $r_{m,n}$ on $Y \times X$ such that for each m and n, for ν-almost all y,

$$\langle b_y(x) e_m, e_n \rangle = r_{n,m}(y,x) \tag{32}$$

for α-almost all x. There is a set $N \subseteq Y$ of ν-measure zero such that for each $y \in Y - N$ and all m, n, (32) holds for α-almost all x. Since $|b_y(x)| \leq 1$, we deduce from (32) and the Fubini theorem that for $(\nu \times \alpha)$-almost all y, x,

$$\sum_n |r_{n,m}(y,x)|^2 < \infty \tag{33}$$

for each m, and

$$\sum_n \left| \sum_{m=1}^{k} r_{n,m}(y,x)c_m \right|^2 \leq |c_1|^2 + \cdots + |c_k|^2 \tag{34}$$

for all k and all k-tuples of complex numbers $(c_1, \cdots, c_k)$ whose real and imaginary parts are rational. Let $Z \subseteq Y \times X$ be the Borel set of all (y,x) such that (33) and (34) are satisfied for the indicated ranges of m, k, c_i. The complement of Z is obviously $\nu \times \alpha$-null. Let $r^*_{n,m}$ be defined by

$$r^*_{n,m}(y,x) = \begin{cases} r_{n,m}(y,x) & (y,x) \in Z, \\ 0 & (y,x) \notin Z. \end{cases} \tag{35}$$

Then the $r^*_{n,m}$ are Borel functions and the inequalities (33) and (34) are satisfied by the $r^*_{n,m}$ for all y, x. Hence for each y, x there exists an operator $b(y,x)$ such that $|b(y,x)| \leq 1$ and

$$\langle b(y,x)e_m, e_n \rangle = r^*_{n,m}(y,x). \tag{36}$$

Equation (36) shows that $(y,x) \to b(y,x)$ is a Borel map of $Y \times X$ into S. Now, $r_{n,m} = r^*_{n,m}$ $(\nu \times \alpha)$-almost everywhere on $Y \times X$, so that we easily deduce from (32) that for ν-almost all y,

$$\langle b_y(x)e_m, e_n \rangle = \langle b(y,x)e_m, e_n \rangle$$

for α-almost all x, i.e., for ν-almost all y,

$$B_y = b(y, \cdot)^{\sim}.$$

If b' is another bounded Borel map of $Y \times X$ into S satisfying (31) for ν-almost all y, it follows from lemma 6.4 and Fubini's theorem that $b(y,x) = b'(y,x)$ for $(\nu \times \alpha)$-almost all y, x. For the last statement, note that if B_y is unitary for all y, $Z = \{(y,x) : b(y,x) \text{ is unitary}\}$ is a Borel set whose complement is of $(\nu \times \alpha)$-measure zero. All we have to do is to redefine $b(y,x)$ to be 1 for $(y,x) \notin Z$. The proof of the lemma is complete.

3. FROM COCYCLES TO SYSTEMS OF IMPRIMITIVITY

We remarked in our example 2 of Section 1 on the connection between the cocycle identities (35) of Chapter V and the identities (6). The main aim of this section is to show that associated with a cocycle taking values in a unitary group, there is a system of imprimitivity, and that this system is determined up to equivalence by the cohomology class of the cocycle. This is precisely the reason for our detailed study of cocycles in the previous chapter. We begin with an elementary lemma.

Lemma 6.6. *Let G be a lcsc group and $\mathscr{H}$ a separable Hilbert space. Suppose that $L(g \to L_g)$ is a mapping of G into the set of bounded operators in $\mathscr{H}$ such that* (i) *for all $f, f' \in \mathscr{H}$, $g \to \langle L_g f, f' \rangle$ is a measurable function,* (ii) *L_g is unitary for almost all g, and* (iii) *$L_{g_1 g_2} = L_{g_1} L_{g_2}$ for almost all $(g_1, g_2) \in G \times G$. Then there exists exactly one representation U of G in $\mathscr{H}$ such that $L_g = U_g$ for almost all g.*

Proof. Let μ be a left Haar measure on G and $\mathscr{L}^1$ the Banach space of μ-summable functions. For $f_1, f_2 \in \mathscr{L}^1$ we introduce their convolution

$$(f_1 * f_2)(g) = \int_G f_1(t) f_2(t^{-1} g) d\mu(t).$$

$f_1 * f_2$ is almost everywhere finite and defines an element of $\mathscr{L}^1$. Define the operator L_f, for $f \in \mathscr{L}^1$, by setting, for $u, v \in \mathscr{H}$,

$$\langle L_f u, v \rangle = \int_G f(g) \langle L_g u, v \rangle d\mu(g). \tag{37}$$

$f \to L_f$ is linear and $\|L_f\| \leq \|f\|_1$. A routine use of the Fubini theorem, coupled with (iii) of the lemma, gives us the equation $L_{f_1 * f_2} = L_{f_1} L_{f_2}$. Moreover, the sum of the ranges of the L_f is dense in $\mathscr{H}$; for, if $v \in \mathscr{H}$ and

$$\int_G f(g) \langle L_g u, v \rangle d\mu(g) = 0$$

for all $u \in \mathscr{H}$ and $f \in \mathscr{L}^1$, then, for each $u \in \mathscr{H}$, $\langle L_g u, v \rangle = 0$ for almost all g. Taking a countable dense set of u's, we obtain the equation $L_g^* v = 0$ for almost all g. Hence $v = 0$.

We now apply a known theorem (Loomis [2], p. 128) and obtain a unique homomorphism $g \to U_g$ of G into the group of invertible operators in $\mathscr{H}$ such that $\|U_g\| \leq 1$ for each g, $g \to U_g$ is Borel, and for each $f \in \mathscr{L}^1$,

$$\int_G f(g) \langle U_g u, v \rangle d\mu(g) = L_f \qquad (u, v \in \mathscr{H}).$$

This implies rather easily that $U_g = L_g$ for almost all g. U_g is therefore unitary for almost all g. Now, the set of g, for which U_g is unitary, is a subgroup. Hence U_g is unitary for all g, i.e., U is a representation. The uniqueness of U is obvious.

Let X be a standard Borel G-space. We choose an invariant measure class $\mathscr{C}$ on X and fix it throughout the rest of this section. Let α be a measure in this class. We introduce a separable Hilbert space $\mathscr{K}$ and denote its unitary group by M. Let $\mathscr{H} = \mathscr{L}^2(X,\mathscr{K},\alpha)$. By a cocycle we mean a (G,X,M)-cocycle relative to α, i.e., a Borel map $(g,x) \to \varphi(g,x)$ of $G \times X$ into M such that the identities (35) of Chapter V are satisfied. The null sets on X are those of $\mathscr{C}$, and on G the usual ones.

Theorem 6.7. *Let the notation be as above. For each $g \in G$, let r_g be a positive Borel function which is a version of $d\alpha/d\alpha^{(g^{-1})}$. Suppose that φ is a (G,X,M)-cocycle relative to α. Then, there exists a unique system of imprimitivity (U,P) such that*

$$P_E f = \chi_E f, \tag{38}$$

and for almost all g,

$$(U_g f)(x) = \{r_g(g^{-1}\cdot x)\}^{1/2}\varphi(g,g^{-1}\cdot x)f(g^{-1}\cdot x) \tag{39}$$

for each $f \in \mathscr{H}$ for almost all x. Moreover, the equivalence class of (U,P) depends only on $\mathscr{C}$ and γ_φ, where γ_φ is the cohomology class of φ.

Proof. Equation (38) defines P_E and $P(E \to P_E)$ is a projection valued measure in $\mathscr{H}$. Let us now define, for each $g \in G$ and $f \in \mathscr{H}$, the Borel function $L_g f$ on X by

$$(L_g f)(x) = \{r_g(g^{-1}\cdot x)\}^{1/2}\varphi(g,g^{-1}\cdot x)f(g^{-1}\cdot x). \tag{40}$$

Since φ takes values in M, we have:

$$|(L_g f)(x)|^2 = r_g(g^{-1}\cdot x)|f(g^{-1}\cdot x)|^2,$$

from which we conclude that $L_g f \in \mathscr{H}$ and $\|L_g f\| = \|f\|$. L_g is thus an isometry. It is actually unitary because its range is the whole of $\mathscr{H}$. In fact, if $f' \in \mathscr{H}$ and $\langle L_g f : f'\rangle = 0$ for all $f \in \mathscr{H}$, a brief calculation shows that $\varphi(g,x)^* \cdot f'(g\cdot x) = 0$ for almost all x which gives us the relation $f' = 0$. Next, φ being a cocycle,

$$\varphi(g_1g_2,x) = \varphi(g_1,g_2\cdot x)\varphi(g_2,x)$$

for almost all $(g_1,g_2,x) \in G \times G \times X$. From this we conclude by a routine calculation that $L_{g_1g_2} = L_{g_1}L_{g_2}$ for almost all g_1, g_2 (for the r_g we use identities (20) of Chapter V). We now verify that $g \to L_g$ is a measurable map. We choose a positive Borel function r on $G \times X$ such that, for almost all g, $r(g,x) = r_g(x)$ for almost all x. Then, for $f, f' \in \mathscr{H}$,

$$\langle L_g f : f'\rangle = \int_G \{r(g,g^{-1}\cdot x)\}^{1/2}\langle\varphi(g,g^{-1}\cdot x)f(g^{-1}\cdot x),f'(x)\rangle d\alpha(x)$$

for almost all g. Since the integrand on the right is a Borel function on $G \times X$, the integral is a Borel function on G and hence $g \to \langle L_g f : f'\rangle$ is a measurable function on G. We are therefore in a position to apply lemma 6.6. Thus there exists a unique representation U of G in $\mathscr{H}$ such that $U_g = L_g$ for almost all g.

We calculate from (40) that for each $g \in G$, $L_g P_E = P_{g \cdot E} L_g$ for all Borel sets $E \subseteq X$. Therefore, if N is a Borel set of measure zero such that $U_g = L_g$ for $g \in G - N$, then for each $g \in G - N$,

$$U_g P_E U_g{}^{-1} = P_{g \cdot E} \tag{41}$$

for all Borel sets E. Now, as U is a representation, the set of all $g \in G$ for which (41) holds for all Borel E is a subgroup of G. As the complement of this subgroup has measure zero, we conclude that (41) must be valid for all $g \in G$ and all Borel sets E. (U,P) is thus a system of imprimitivity based on X. Its uniqueness is obvious.

To show that the equivalence class of (U,P) depends only on the measure class of α and the cohomology class of φ, we proceed in two steps. First, we replace α by α' in the same measure class, keeping φ fixed. In this case, let $a > 0$ be a Borel function such that $a = d\alpha'/d\alpha$. Let (U',P') be the system of imprimitivity associated with α' and φ. Then

$$W : f \to a^{-1/2} f \tag{42}$$

is a unitary isomorphism of $\mathscr{L}^2(X,\mathscr{K},\alpha)$ onto $\mathscr{L}^2(X,\mathscr{K},\alpha')$ and the formula (39) shows that

$$P_E{}' = W P_E W^{-1}$$

for all E, and for almost all g,

$$U_g{}' = W U_g W^{-1}.$$

Once again, as the set of all g such that $U_g{}' = W U_g W^{-1}$ is a subgroup, these equations are valid for all g. Thus (U,P) and (U',P') are equivalent.

It remains to examine what happens when we fix α but change φ. Let φ' be a cocycle defining the same cohomology class as φ. Let (U',P') be the system of imprimitivity associated with α, φ'. Both (U',P') and (U,P) act in $\mathscr{H} = \mathscr{L}^2(X,\mathscr{K},\alpha)$. Since $\varphi' \sim \varphi$, there exists a Borel map b,

$$b : x \to b(x) \tag{43}$$

of X into M such that

$$\varphi'(g,x) = b(g \cdot x)\varphi(g,x)b(x)^{-1}$$

for almost all $(g,x) \in G \times X$. For any $f \in \mathscr{H}$,

$$|f(x)|^2 = |b(x)f(x)|^2,$$

and hence there exists a unitary operator B of $\mathscr{L}^2(X,\mathscr{K},\alpha)$ such that (cf. Section 2)

$$B = b^{\sim}. \tag{44}$$

From (44) we obtain

$$BP_E = P_E B \tag{45}$$

for all Borel sets E. Next we compute BU_gB^{-1}. We have, from (39), for almost all $g \in G$,

$$\begin{aligned}(BU_gB^{-1}f)(x) &= \{r_g(g^{-1}\cdot x)\}^{1/2}b(x)\varphi(g,g^{-1}\cdot x)b(g^{-1}\cdot x)^{-1}f(g^{-1}\cdot x)\\ &= \{r_g(g^{-1}\cdot x)\}^{1/2}\varphi'(g,g^{-1}\cdot x)f(g^{-1}\cdot x)\\ &= (U_g'f)(x)\end{aligned}$$

for each $f \in \mathscr{H}$ for almost all x. Therefore

$$U_g' = BU_gB^{-1} \tag{46}$$

for almost all g. Using our usual arguments, we conclude that (46) is valid for all g. Equations (46) and (45) show that (U,P) and (U',P') are equivalent. This completes the proof of the theorem.

Corollary 6.8. *If φ is a strict cocycle or even if we assume that for each $(g_1,g_2) \in G \times G$,*

$$\varphi(g_1g_2,x) = \varphi(g_1,g_2\cdot x)\varphi(g_2,x)$$

for almost all x, then for each $g \in G$ and each f

$$(U_gf)(x) = \{r_g(g^{-1}\cdot x)\}^{1/2}\varphi(g,g^{-1}\cdot x)f(g^{-1}\cdot x)$$

for almost all x.

In this case, $g \to L_g$ itself is a homomorphism so that $L_g = U_g$ for all g.

Given φ, α as above, the system of imprimitivity (U,P) constructed in theorem 6.7 is said to be *associated* with φ and α.

4. PROJECTION VALUED MEASURES

In this book we have come across projection valued measures from time to time, but such uses as we have made of them have been quite superficial. The aim of this section is to describe some of the deeper aspects of the theory of projection valued measures. These results are all known and form the Hahn–Hellinger multiplicity theory. We recommend the reader to the accounts of Stone [2], Nakano [1], and Halmos [2]. We shall confine ourselves to a concise description of the main features of this theory.

The main problem in this theory is that of determining when two projection valued measures are equivalent. More precisely, let X be a standard Borel space, $\mathscr{H}^i$ a separable Hilbert space, and P^i a projection valued measure in $\mathscr{H}^i$ based on X. We say P^1 and P^2 are equivalent, $P^1 \simeq P^2$ in symbols, when there is a unitary isomorphism W of $\mathscr{H}^1$ onto $\mathscr{H}^2$ such that $P_E{}^2 = WP_E{}^1W^{-1}$ for all E. The main problem is, given X, to construct *canonical forms* for the projection valued measures based on X.

We begin by choosing a σ-finite measure α on X and a separable Hilbert space $\mathscr{K}$. Let $\mathscr{H} = \mathscr{L}^2(X,\mathscr{K},\alpha)$ and P be the projection valued measure $E \rightarrow P_E$, where

$$P_E f = \chi_E f \qquad (f \in \mathscr{H}). \tag{47}$$

We write $P = P(\mathscr{K},\alpha)$. The first result of the theory is that $P(\mathscr{K},\alpha) \simeq P(\mathscr{K}',\alpha')$ if and only if (i) α and α' define the same measure class, and (ii) $\dim \mathscr{K} = \dim \mathscr{K}'$. If these conditions are satisfied, if $\mathscr{K} = \mathscr{K}'$, and if a is a positive Borel function which is a version of $d\alpha'/d\alpha$, the map

$$W : f \rightarrow a^{-1/2} f$$

is a unitary isomorphism of $\mathscr{L}^2(X,\mathscr{K},\alpha)$ onto $\mathscr{L}^2(X,\mathscr{K},\alpha')$ such that $P(\mathscr{K},\alpha') = W P(\mathscr{K},\alpha) W^{-1}$.

This result motivates us to introduce the following definition. A projection valued measure P (acting in a separable Hilbert space) is said to be *homogeneous* if it is equivalent to $P(\mathscr{K},\alpha)$ for suitable $\mathscr{K}$ and α. The definition (47) of $P(\mathscr{K},\alpha)$ shows that for a Borel set $E \subseteq X$, $\alpha(E) = 0$ if and only if $P_E = 0$. The measure class of α is thus determined by P in a very direct fashion. It is called the *measure class of* P. The dimension of $\mathscr{K}$ is said to be the *multiplicity of* P. Given any measure class $\mathscr{C}$ on X and any integer n $(1 \leq n \leq \infty)$, there exists, up to equivalence, a unique homogeneous projection valued measure P with $\mathscr{C}$ as the measure class of P and n its multiplicity.

On the other hand, given a projection valued measure P on X, it is a deeper matter to determine when it is homogeneous. A vector $u \in \mathscr{H}$ is said to be *cyclic* for P if the collection $\{P_E u\}$ spans a dense linear manifold of $\mathscr{H}$. It can be shown that if there is a cyclic vector for P, P is homogeneous and of multiplicity one and that, conversely, homogeneous projection valued measures of multiplicity one do have this geometric property, i.e., existence of a cyclic vector. The case of arbitrary multiplicity can be handled along similar, but somewhat more complicated, lines.

Now we pass on to the general case. For each $n = \infty, 1, 2, \cdots$ we choose a separable Hilbert space $\mathscr{K}_n$ of dimension n and fix it throughout the discussion. Choose a sequence of σ-finite measures $\alpha_\infty, \alpha_1, \alpha_2, \cdots$ on X. For any σ-finite measure α on X let us write $\mathscr{H}_{n,\alpha} = \mathscr{L}^2(X,\mathscr{K}_n,\alpha)$ and $P^{n,\alpha} = P(\mathscr{K}_n,\alpha)$. *Suppose that the* α_n *are so chosen that for* $m \neq n$, α_m *and* α_n *are mutually singular.*† Let

$$\mathscr{H} = \bigoplus_n \mathscr{H}_{n,\alpha_n} \tag{48}$$

be the direct sum of the $\mathscr{H}_{n,\alpha_n}$, and for $E \subseteq X$ let P_E be defined by

$$P_E : (f_\infty, f_1, f_2, \cdots) \rightarrow (P_E^{\infty,\alpha_\infty} f_\infty, P_E^{1,\alpha_1} f_1, \cdots) \tag{49}$$

---

† σ-finite measures α and β on X are called mutually singular if we can write $X = A \cup B$, where A and B are Borel sets with $A \cap B = \varnothing$, $\beta(A) = \alpha(B) = 0$.

(when $\alpha_n = 0$ for some n, $\mathscr{H}_{n,\alpha_n} = 0$). Then P is a projection valued measure. We write $P = P(\{\mathscr{K}_n\}, \{\alpha_n\})$. The main theorem of spectral multiplicity theory asserts that every projection valued measure on X in a separable Hilbert space is equivalent to $P(\{\mathscr{K}_n\}, \{\alpha_n\})$ for a suitable sequence $\{\alpha_n\}$. If $\{\alpha_n\}$ and $\{\alpha_n'\}$ are two sequences of mutually singular measures on X, the theorem further asserts that

$$P(\{\mathscr{K}_n\}, \{\alpha_n\}) \simeq P(\{\mathscr{K}_n\}, \{\alpha_n'\})$$

if and only if for each n, α_n and α_n' are in the same measure class.

Given P, we therefore have for each $n = \infty, 1, 2, \cdots$ a canonically determined measure class, say $\mathscr{C}_n$. $\mathscr{C}_n$ is called *the measure class corresponding to the multiplicity* n. For a given n, $\mathscr{C}_n$ may be 0. P is homogeneous if and only if $\mathscr{C}_n = 0$ for all n except a single value n_0. The construction of the $\mathscr{C}_n$ from P is delicate. We do not need this for our purposes. On the other hand, consider the set of all σ-finite measures α with the property that $\alpha(E) = 0$ if and only if E is $\mathscr{C}_n$-null for all n $(1 \leq n \leq \infty)$. There exist such measures α; in fact, if we choose for each n a finite α_n in $\mathscr{C}_n$ such that $\sum_n \alpha_n(X) < \infty$, then $\alpha = \sum_n \alpha_n$ would have the required property. These measures α form a measure class, say $\mathscr{C}$. It is obvious that a set E is $\mathscr{C}$-null if and only if $P_E = 0$. $\mathscr{C}$ is called the *measure class of* P.

5. FROM SYSTEMS OF IMPRIMITIVITY TO COCYCLES

We now examine the problem converse to that studied in Section 3. We resume our concern with a standard Borel G-space X and a system of imprimitivity (U,P) based on X, acting in $\mathscr{H}$. These will be fixed throughout this section. The main result is that if P is homogeneous, one can construct a cocycle φ relative to the measure class $\mathscr{C}$ of P such that (U,P) is equivalent to the system of imprimitivity associated with $\mathscr{C}$ and φ. We shall use the notations and results described in Section 4 concerning projection valued measures. In particular, for each $n(\infty, 1, 2, \cdots)$, $\mathscr{K}_n$ is a fixed separable Hilbert space of dimension n. We write M_n for the unitary group of $\mathscr{K}_n$. The norm and inner product in $\mathscr{K}_n$ are denoted by $|\,.\,|$ and $\langle .,. \rangle$, while the norm and inner product in $\mathscr{H}$ and $\mathscr{H}_{n,\alpha}$ (α a σ-finite measure on X) are denoted by $\|\,.\,\|$ and $\langle . : . \rangle$.

Lemma 6.9. *With the notation described above, the measure classes* $\mathscr{C}_n$ *corresponding to the various multiplicities* n *are all invariant under the action of* G. *In particular, the measure class of* P *is invariant.*

Proof. Choose measures $\alpha_n \in \mathscr{C}_n$. Then, $P \simeq P(\{\mathscr{K}_n\}, \{\alpha_n\})$. We must show that each $\mathscr{C}_n$ is invariant under G. Let $g \in G$ be arbitrary. Let us write $Q_E = P_{g \cdot E}$. Then $Q(E \to Q_E)$ is a projection valued measure in $\mathscr{H}$. The equation

$$U_g P_E U_g^{-1} = P_{g \cdot E}$$

shows that $Q \simeq P$. But it is obvious that $Q \simeq P(\{\mathscr{H}_n\},\{\alpha_n^g\})$ (via the map $f \to f^g$) and that the α_n^g are mutually singular. Hence for each n, α_n and α_n^g define the same measure class. This proves the lemma.

We recall that a measure class $\mathscr{C}$ on X is said to be ergodic if any Borel function f, with the property that $f^g = f$ $\mathscr{C}$-almost everywhere for each $g \in G$, is constant $\mathscr{C}$-almost everywhere.

Lemma 6.10. *If the measure class $\mathscr{C}$ of P is ergodic, then P is homogeneous. If (U,P) is irreducible, $\mathscr{C}$ is ergodic and so P is homogeneous.*

Proof. Choose mutually singular finite measures α_n $(n = \infty, 1, 2, \cdots)$ such that we then have $P \simeq P(\{\mathscr{H}_n\},\{\alpha_n\})$. Let $\mathscr{C}_n$ be the measure class defined by α_n. Since the α_m are mutually singular, we can find mutually disjoint Borel sets E_n $(n = \infty, 1, 2, \cdots)$ such that $X = \bigcup_{1 \leq n \leq \infty} E_n$ and $\alpha_n(X - E_n) = 0$ for all n. Suppose that the measure class $\mathscr{C}$ of P is ergodic, but that for two *distinct* values of n, $\mathscr{C}_n \neq 0$. Let a_n be distinct real numbers and let f be the Borel function which takes the value a_n on E_n. For each n, f is a constant $\mathscr{C}_n$-almost everywhere. Since $\mathscr{C}_n \neq 0$ implies $\alpha_n(E_n) \neq 0$, E_n is not $\mathscr{C}$-null for any n for which $\mathscr{C}_n$ is nonzero. Thus f is not a constant $\mathscr{C}$-almost everywhere. But (by lemma 6.9) as each $\mathscr{C}_n$ is invariant, $f^g = f$, $\mathscr{C}$-almost everywhere for each $g \in G$. This is a contradiction. We must therefore have $\mathscr{C}_n = 0$ for all but a single value of n, i.e., P is homogeneous. The second statement would follow if we show that $\mathscr{C}$ is ergodic whenever (U,P) is irreducible. Suppose now that (U,P) is irreducible but $\mathscr{C}$ is not ergodic. Then there exists a real Borel function f such that f is not a constant $\mathscr{C}$-almost everywhere but, for each $g \in G$, $f^g = f$, $\mathscr{C}$-almost everywhere. Let c be a real number such that, if $E_0 = \{x : f(x) < c\}$, then neither E_0 nor $X - E_0$ is $\mathscr{C}$-null. If $g \in G$, the fact that $f(g^{-1} \cdot x) = f(x) = f(g \cdot x)$ for $\mathscr{C}$-almost all x implies that the set $(g \cdot E_0 - E_0) \cup (E_0 - g \cdot E_0)$ is $\mathscr{C}$-null. Hence $P_{g \cdot E_0} = P_{E_0}$. Since this is true for all g, (6) shows that P_{E_0} commutes with all U_g. Since P_{E_0} obviously commutes with all P_E, the irreducibility of (U,P) implies that either P_{E_0} or $P_{X - E_0}$ is 0, i.e., either E_0 or $X - E_0$ is null, a contradiction.

Theorem 6.11. *Let (U,P) be a system of imprimitivity acting in $\mathscr{H}$ such that P is homogeneous and of multiplicity n $(1 \leq n \leq \infty)$. Let α be a σ-finite measure in the measure class $\mathscr{C}_P$ of P. Then (U,P) is equivalent to a system $(U',P^{n,\alpha})$ acting in $\mathscr{H}_{n,\alpha}$. Further, for each cohomology class γ of (G,X,M_n)-cocycles relative to $\mathscr{C}_P$, let $\Sigma(\gamma)$ denote the equivalence class of systems of imprimitivity acting in $\mathscr{H}_{n,\alpha}$ which are associated to γ and $\mathscr{C}_P$. Then*

$$\gamma \rightleftarrows \Sigma(\gamma) \tag{50}$$

is a one-one correspondence between the set of all cohomology classes of (G,X,M_n)*-cocycles relative to* $\mathscr{C}_P$ *and the set of all equivalence classes of systems of imprimitivity of the form* $(U',P^{n,\alpha})$.

Proof. There exists a unitary isomorphism W of $\mathscr{H}$ onto $\mathscr{H}_{n.\alpha}$ such that

$$P_E^{n,\alpha} = WP_EW^{-1}$$

for all Borel sets E. We set $U_g' = WU_gW^{-1}$. Then $(U,P) \simeq (U',P^{n,\alpha})$.

Consider now a system $(U',P^{n,\alpha})$. Let us choose, for each g in G, a positive Borel function r_g such that r_g is a version of $d\alpha/d\alpha^{(g^{-1})}$. For simplicity of notation we write $\mathscr{H}' = \mathscr{H}_{n,\alpha}$, $P' = P^{n,\alpha}$. Let (V,P') be the Koopman system of imprimitivity, i.e., for each $g \in G$ and $f \in \mathscr{H}'$,

$$(V_gf)(x) = \{r_g(g^{-1}\cdot x)\}^{1/2} \qquad f(g^{-1}\cdot x). \tag{51}$$

Write

$$W_g = V_g^{-1}U_g'. \tag{52}$$

Since both (U',P') and (V,P') are systems of imprimitivity, it follows that W_g commutes with P_E for all E. Further, $g \to W_g$ is a Borel map of G into the unitary group of $\mathscr{H}'$. We now use lemmas 6.4 and 6.5. We select, for each $g \in G$, a Borel map w_g of X onto M_n such that

$$W_g = w_g^{\sim}, \tag{53}$$

and a Borel map φ' of $G \times X$ into M_n,

$$\varphi' : g, x \to \varphi'(g,x), \tag{54}$$

such that for almost all g,

$$W_g = \varphi'(g,\cdot)^{\sim}. \tag{55}$$

For $g_1, g_2 \in G$, we obtain from (52),

$$W_{g_1g_2} = (V_{g_2}^{-1}W_{g_1}V_{g_2})W_{g_2}. \tag{56}$$

A quick calculation shows that

$$V_{g_2}^{-1}W_{g_1}V_{g_2} = (w_{g_1}^{(g_2^{-1})})^{\sim}, \tag{57}$$

where, as usual, for $x \in X$,

$$w_{g_1}^{(g_2^{-1})}(x) = w_{g_1}(g_2\cdot x).$$

From (56) and (57) and lemma 6.4, it follows that for each $(g_1,g_2) \in G \times G$,

$$w_{g_1g_2}(x) = w_{g_1}(g_2\cdot x)w_{g_2}(x) \tag{58}$$

for almost all x. Then

$$\varphi'(g_1g_2,x) = \varphi'(g_1,g_2\cdot x)\varphi'(g_2,x) \tag{59}$$

for almost all $(g_1,g_2,x) \in G \times G \times X$. This proves that φ' is a (G,X,M_n)-cocycle relative to α. Moreover, as $U_g' = V_g W_g$, we have, for almost all g, $U_g' = V_g \varphi'(g,\cdot)^\sim$, i.e., for almost all g,

$$(U_g' f)(x) = \{r_g(g^{-1}\cdot x)\}^{1/2} \varphi'(g,g^{-1}\cdot x) f(g^{-1}\cdot x) \tag{60}$$

for $f \in \mathscr{H}'$. In other words, $(U', P^{n,\alpha})$ is the system of imprimitivity associated with φ' and α.

If φ_1' is another cocycle satisfying (60), then, for almost all g, $\varphi_1'(g,\cdot)^\sim = W_g = w_g^\sim$. This shows that $\varphi_1' = \varphi'$ almost everywhere on $G \times X$.

To complete the proof of the theorem, we have only to show that if (U'',P') is a second system of imprimitivity equivalent to (U',P') and if φ'' is a cocycle corresponding to (U'',P'), then φ' and φ'' determine the same cohomology class. Let W be a unitary automorphism of $\mathscr{H}'$ such that

$$\begin{aligned} W P_E' W^{-1} &= P_E', \\ W U_g'' W^{-1} &= U_g', \end{aligned} \tag{61}$$

for all g, E. There exists (lemma 6.4) a Borel map w,

$$w : x \to w(x) \tag{62}$$

of X into M_n such that $W = w^\sim$. But then the equation $W U_g'' W^{-1} = U_g'$, together with the relations between φ'' and U'' as well as φ' and U', show, after a brief calculation, that for almost all $(g,x) \in G \times X$,

$$w(x)\varphi''(g,g^{-1}\cdot x) w(g^{-1}\cdot x)^{-1} = \varphi'(g,g^{-1}\cdot x),$$

i.e.,

$$w(g\cdot x)\varphi''(g,x) w(x)^{-1} = \varphi'(g,x)$$

for almost all $(g,x) \in G \times X$. This shows that $\varphi' \sim \varphi''$. This completes the proof of the theorem.

Remark. It is easy to see that if the cocycle φ' is a coboundary, the corresponding systems of imprimitivity are equivalent to the Koopman systems. Thus, the examples of Section 5 in Chapter V give rise to systems (U,P) which are not equivalent to the Koopman systems.

6. TRANSITIVE SYSTEMS

In this section X continues to be a standard Borel G-space and (U,P) a system of imprimitivity based on X. We shall completely describe all the *transitive* systems. Our success is due to the detailed knowledge of cocycles on transitive G-spaces which was accumulated in Section 5 of Chapter V. We recall that a measure class $\mathscr{C}$ is said to be transitive if there exists $x_0 \in X$ such that $X - G\cdot x_0$ is a $\mathscr{C}$-null set. In this case the orbit $G \cdot x_0$ is uniquely determined by $\mathscr{C}$ and we shall say that $\mathscr{C}$ *lives on* $G \cdot x_0$. A system of imprimitivity (U,P) is said to be *transitive* if the measure class $\mathscr{C}$ of P is such; if X' is the orbit on which $\mathscr{C}$ lives, we say that the system *lives* on X'. Note that if (U,P) is transitive, the measure class of P is ergodic, and hence P is homogeneous (lemma 6.10). If $x \in X$ and $X' = G \cdot x$, X' is an invariant Borel set and is hence a standard transitive Borel G-space in its own right. If α' is a σ-finite quasi-invariant measure on X', the measure $\alpha : E \to \alpha'(E \cap X')$ defines a measure class on X which is invariant and which lives on X'. Obviously it is the only measure class living on X' with this property.

Let $x_0 \in X$ and $X' = G \cdot x_0$. Suppose that $G_{x_0} = G_0$ is the stability subgroup at x_0. Let m be a representation of G_0 in the Hilbert space $\mathscr{K}_n$ (which we have chosen once for all) of dimension n $(1 \leq n \leq \infty)$. Then m is a Borel homomorphism of G_0 into M_n. From the results of Chapter V we know that there exists a strict (G,X',M_n)-cocycle φ_0, which defines the homomorphism m of G_0 at x_0. We extend φ_0 to a strict (G,X,M_n)-cocycle φ, for example, by putting $\varphi(g,x) = 1$ for $x \notin X'$. We choose a σ-finite measure α which is quasi invariant and lives on X'. We can then construct the system of imprimitivity in $\mathscr{K}_{n,\alpha}$, associated with φ and α. We shall call such systems, obtained by varying α, φ (but α living on X') systems *induced by the representation m of G_0*. It is obvious that all such systems are mutually equivalent, homogeneous, transitive, and live on X'. We now prove the basic theorem of this section.

Theorem 6.12. *Let X be a standard Borel G-space and let (U,P) be a transitive system of imprimitivity based on X such that* (i) *P is of multiplicity n, and* (ii) *the measure class of P lives on the orbit $X' = G \cdot x_0$. Then there exists a representation m of the stability group G_0 at x_0 in $\mathscr{K}_n$ such that (U,P) is equivalent to a system induced by m. The equivalence class of m is determined uniquely by the equivalence class of (U,P) and gives rise to a one-one correspondence $\{(U,P)\} \to \{m\}$ of the set of all equivalence classes of transitive systems living on X' onto the set of all equivalence classes of representations of G_0. Moreover, the commuting ring of (U,P) is isomorphic to the commuting ring of m. In particular, (U,P) is irreducible, or a direct sum of irreducible systems, if and only if m is so.*

Proof. If m_1 and m_2 are inequivalent representations of G_0, theorem 5.27 shows that the (G,X',M_n)-cocycles which correspond to them define distinct cohomology classes. Therefore we can conclude from theorem 6.11 that the systems of imprimitivity induced by m_1 and m_2 are inequivalent. By the same token, if m_1 and m_2 are equivalent, so are the systems that are induced by m_1 and m_2.

Next, let (U,P) be a transitive system which lives on X'. Let α be a σ-finite measure belonging to the measure class of P. $\alpha(X - X') = 0$

obviously. Theorem 6.11 now tells us that there exists some n and a (G,X,M_n)-cocycle φ relative to α such that (U,P) is equivalent to the system associated with φ and α. Now, the restriction of φ to $G\times X'$ is a (G,X',M_n)-cocycle on the transitive space X' and hence by lemma 5.26 there exists a strict cocycle φ' on $G\times X'$ which defines the same cohomology class as φ (restricted to $G\times X'$). If m is the homomorphism of G_0 which is defined by φ', it is clear that (U,P) is equivalent to a system induced by m.

The only thing that remains to be proved is the statement about the commuting rings. We assume (as we may) that $X=X'$. Let φ be a strict (G,X,M_n)-cocycle defining the homomorphism m at x_0 of G_0. We choose a quasi-invariant σ-finite measure α on X and write $\mathscr{H}=\mathscr{H}_{n,\alpha}$. In our usual notation, for each $g\in G$ and Borel $E\subseteq X$,

$$\begin{aligned} P_E f &= \chi_E f, \\ (U_g f)(x) &= \{r_g(g^{-1}x)\}^{1/2}\varphi(g,g^{-1}x)f(g^{-1}x). \end{aligned} \tag{63}$$

We use the results and notation of Section 5, Chapter V. We select a Borel map $d(g\to d(g))$ of G into M_n such that $d(gh)=d(g)m(h)$ for all $(g,h)\in G\times G_0$, $d(e)=1$, and $\varphi^0(g,g')=d(gg')d(g')^{-1}$ for all $(g,g')\in G\times G$ $(\varphi^0(g,g')=\varphi(g,g'\cdot x_0))$.

Suppose that T is an operator in $\mathscr{H}$ which commutes with all P_E and all U_g. By lemma 6.4 we can choose a Borel map t,

$$t : x\to t(x) \tag{64}$$

of X into the Borel space S_n of all operators in $\mathscr{H}_n$ such that $|t(x)|\leq\|T\|$ for all x and

$$T = t^{\sim}. \tag{65}$$

If we use the fact that T commutes with all U_g, we obtain, after a brief calculation, the identity

$$\varphi(g,g^{-1}\cdot x)t(g^{-1}\cdot x) = t(x)\varphi(g,g^{-1}\cdot x) \tag{66}$$

for each g for almost all x, i.e., for each $g\in G$

$$\varphi(g,x)t(x) = t(g\cdot x)\varphi(g,x) \tag{67}$$

for almost all $x\in X$. From (67) we get, on writing $t^0(g)=t(g\cdot x_0)$,

$$\varphi^0(g,g')t^0(g') = t^0(gg')\varphi^0(g,g')$$

for almost all $(g,g')\in G\times G$. Substituting the expression for φ^0 in terms of d and simplifying, we conclude that there exists a bounded operator τ in $\mathscr{H}_n$ such that

$$t^0(g) = d(g)\tau d(g)^{-1} \tag{68}$$

for almost all g. Replacing g by gh ($h \in G_0$) and simplifying, we obtain

$$m(h)\tau m(h)^{-1} = \tau. \tag{69}$$

This shows that τ is in the commuting ring of m. On the other hand, if we start from a τ in the commuting ring of m, the map $g \to d(g)\tau d(g^{-1})$ is a bounded Borel map of X into S_n which is constant on the left G_0-cosets. Hence there is a bounded Borel map t from X into S_n such that t^0 is given by (68) for all g. Routine calculation leads us from (68) to (67) to (66). At this stage, we know that $T = t^\sim$ is in the commuting ring of (U,P).

The correspondence $T \rightleftarrows \tau$ is obviously one-one and is an algebra isomorphism as is easily seen from the formulas (68) and (65) connecting T and τ. Moreover, as each $d(g)$ is unitary, (68) also implies that

$$t^0(g)^* = d(g)\tau^* d(g)^{-1}, \tag{70}$$

so that $T \rightleftarrows \tau$ is adjoint-preserving. Since the commuting ring of a representation (or a system of imprimitivity) determines completely the circumstances under which it is irreducible or a direct sum of irreducibles, the proof of the theorem is complete.

Corollary 6.13. *Let X be a standard transitive Borel G-space, $x_0 \in X$, and G_0 the stability group at x_0. Then every system of imprimitivity based on X is homogeneous and its measure class is the unique invariant measure class on X. There is a canonical one-one correspondence between the equivalence classes of systems based on X and equivalence classes of representations of G_0. This correspondence preserves irreducibility as well as the property of being a direct sum of irreducibles.*

Remark. Theorem 6.12 gives a complete analysis of the transitive homogeneous systems of imprimitivity based on X. It is natural to ask what happens when we give up the assumption of transitivity. According to lemma 6.10 if the measure class $\mathscr{C}$ of P is ergodic, then P is homogeneous. Theorem 6.11 then tells us that we have only to study the cohomology classes relative to $\mathscr{C}$. However, the examples of Section 5 in Chapter V show that the cohomology classes relative to an ergodic invariant measure class are incredibly more complicated than in the transitive case. In view of this it has not been possible to carry out the analysis of ergodic systems of imprimitivity further than that of theorem 6.11.

We note here that if there is any ergodic invariant measure class $\mathscr{C}$ on X, then one can construct *irreducible* systems of imprimitivity (U,P) based on X such that $\mathscr{C}$ is the measure class of P. In fact, let α be a σ-finite measure in $\mathscr{C}$ and let φ be any (G,X,M_1)-cocycle relative to $\mathscr{C}$. Let $\mathscr{H} = \mathscr{H}_{1,\alpha}$ and let (U,P) be the system of imprimitivity associated with φ and α. We may take $\mathscr{K}_1$ to be the complex number space so that $\mathscr{H} = \mathscr{L}^2(X,\alpha)$. We claim that (U,P) is actually irreducible. In fact, let T be an operator commuting with (U,P). By our lemma 9.4 there exists a *complex valued*, bounded Borel function t such that

$$(Tf)(x) = t(x)f(x)$$

for all $f \in \mathscr{H}$. The condition that T commutes with all U_g gives us (67). But, as $t(x)$ is a complex number, it commutes with $\varphi(g,x)$; thus for each $g \in G$,

$$t(g:x) = t(x)$$

for almost all x. The ergodicity of $\mathscr{C}$ now implies that t is a constant almost everywhere. T is thus a multiple of the identity showing that (U,P) is irreducible. In other words, the limitations imposed by the assumption of transitivity are quite real.

A useful direction of investigation would obviously be that of determining when every ergodic invariant measure class is transitive. We formulate two lemmas, each of which gives sufficient conditions for this to happen.

Lemma 6.14. *Let X be a standard Borel G-space. Suppose that there exists a Borel set $D \subseteq X$ which intersects each orbit in exactly one point. Then, every ergodic invariant class on X is transitive.*

Proof. The proof requires the use of the deep result (cf. remark following corollary 5.2) on the universal measurability of the ranges of Borel maps. Let α be a *finite* measure of an arbitrary ergodic invariant measure class. We normalize α so that $\alpha(X) = 1$. If E is any Borel set $\subseteq D$, the set $G \cdot E$, being the image of the standard space $G \times E$ under the Borel map $g, x \rightarrow g \cdot x$, is an α-measurable subset of X. We define

$$\alpha^{\sim}(E) = \alpha(G \cdot E). \tag{71}$$

Now the fact that D meets each orbit exactly once implies that, for disjoint E, the sets $G \cdot E$ are also disjoint. Hence, $\alpha^{\sim}$, as defined by (71), is a measure on the Borel space defined by D and $\alpha^{\sim}(D) = 1$. We claim that for any Borel set $E \subseteq D$, $\alpha^{\sim}(E)$ is either 0 or 1. If this were not so, we can choose a Borel set $E \subseteq D$ such that $0 < \alpha(G \cdot E) < 1$ and the characteristic function χ of the set $G \cdot E$ would be an invariant α-measurable function which is not a constant almost everywhere. Let χ' be a Borel function on X which is α-almost everywhere equal to χ. Then, for each g, $\chi'(g \cdot x) = \chi'(x)$ for α-almost all x. This would then contradict the ergodicity of the measure α.

Thus $\alpha^{\sim}$ is a *two-valued* measure on D. Now, choose a sequence $\{E_n\}$ of Borel subsets of D which *separate* the points of D. By replacing E_n by

$D-E_n$ whenever necessary, we may assume that $\alpha^{\sim}(E_n)=1$ for all n. Put

$$E_0 = \bigcap_{n=1}^{\infty} E_n; \tag{72}$$

$\alpha^{\sim}(E_0)=1$. Since the sequence $\{E_n\}$ separates the points of D, E_0 must consist of a single point, say x_0. But then $\alpha(G\cdot x_0)=1$, showing that α is transitive.

A variation of the argument given just now leads to:

Lemma 6.15. *Suppose that there exists a sequence $\{f_n\}$ of invariant real valued Borel functions on X which separate the orbits of X, i.e., two points x, y of X lie on the same orbit if and only if $f_n(x)=f_n(y)$ for all n. Then every ergodic invariant measure class on X is transitive.*

Proof. Let α be a measure such that $\alpha(X)=1$ and let α be quasi invariant and ergodic. Let $\mathscr{B}'$ be the smallest Borel structure on X which makes all the f_n Borel. Every set in $\mathscr{B}'$ is a Borel subset of X which is also invariant. Hence each such set has α-measure either 0 or 1. Let $\{E_n\}$ be a sequence of sets which generates $\mathscr{B}'$ such that $\alpha(E_n)=1$ for all n. If we put

$$E_0 = \bigcap_{n=1}^{\infty} E_n, \tag{73}$$

then $\alpha(E_0)=1$. Now the class of Borel sets which either contain E_0 or are disjoint from it is a σ-algebra containing all the E_n. Hence every set in $\mathscr{B}'$ has this property. This implies that E_0 is a single orbit; for, if $x_0 \in E_0$ and $X'=G\cdot x_0$, the fact that the f_n separate the orbits implies that

$$X' = \bigcap_{n} \{x : f_n(x) = f_n(x_0)\}, \tag{74}$$

and consequently that $X' \in \mathscr{B}'$; since $X' \cap E_0 \neq \varnothing$, $E_0 \subseteq X'$, so that $X'=E_0$.

Corollary 6.16. *If G is compact, then every ergodic invariant measure class on X is transitive.*

Proof. Using theorem 5.7 it is easy to come down to the case when X is a compact metric space and G acts *continuously* on X. We shall verify that the condition of lemma 6.15 is satisfied. By the Stone-Weierstrass theorem, $C(X)$ is separable. Let $\{u_n\}$ be a countable set dense in $C(X)$, and for $u \in C(X)$ let

$$\bar{u}(x) = \int_G u(g\cdot x)d\mu(g), \tag{75}$$

where μ is the (normalized) Haar measure on G. $\bar{u}$ is continuous on X. Clearly each $\bar{u}$ is invariant. We claim that $\{\bar{u}_n\}$ separates the orbits in X.

Let $x, y \in X$ such that $\bar{u}_n(x) = \bar{u}_n(y)$ for all n. Since $\{u_n\}$ is dense in $C(X)$, we deduce from (75) that $\bar{u}(x) = \bar{u}(y)$ for all $u \in C(X)$. Suppose now x and y were not in the same orbit. Then $G \cdot x$ and $G \cdot y$ are *closed* disjoint subsets of X, and hence there is a $u \in C(X)$ with $0 \leq u \leq 1$ such that $u = 0$ on $G \cdot x$ and $= 1$ on $G \cdot y$. It is obvious that $\bar{u} = 0$ on $G \cdot x$ and $= 1$ on $G \cdot y$. This contradiction proves the corollary.

If every ergodic invariant measure class on X is transitive, we shall say that X has a *smooth orbit structure.*

7. EXAMPLES AND REMARKS

We shall now discuss a number of examples. The simplest example of a transitive space is G itself. If (U,P) is a system based on G, it is transitive and its measure class is the invariant measure class of G. We now apply the results of Section 6. The stability group at $e \in G$ is $\{e\}$ whose only representations are trivial. Any trivial representation is a direct sum of trivial one-dimensional representations. We thus obtain

Theorem 6.17. *Let G act on itself by left translations. Let* $\mathscr{H} = \mathscr{L}^2(\mu)$, *where* μ *is a left Haar measure on G, and let P and U be defined by:*

$$P_E f = \chi_E \cdot f, \qquad (U_g f)(x) = f(g^{-1}x). \tag{76}$$

Then (U,P) *is an irreducible system of imprimitivity based on G. Any irreducible system based on G is equivalent to this one. An arbitrary system of imprimitivity is a direct sum of irreducible systems.*

Rotations in the Plane. Let $X = \mathbf{C}$, the complex plane, and let G be the multiplicative group of complex numbers of modulus 1. We put, for $z \in G$ and $x \in \mathbf{C}$,

$$z \cdot x = zx.$$

X is a Borel G-space. For each $d \geq 0$,

$$E_d = \{x : |x| = d\}$$

is an orbit and these are all the orbits. Since G is compact, all ergodic measure classes are transitive. For $d > 0$, the stability groups associated with E_d are trivial and hence, corresponding to each $d > 0$, there exists, up to equivalence, exactly one irreducible system living on E_d. We shall now describe this. Note that E_1 has an invariant measure, say σ, such that $\sigma(E_1) = 1$. If we define the measure σ_d by

$$\int_{E_d} f(x) d\sigma_d(x) = \int_{E_1} f(dx) d\sigma(x),$$

then $\sigma_d(E_d)=1$ and it is the unique invariant measure living on E_d with this property. Let $\mathscr{H}_d$ be the Hilbert space $\mathscr{L}^2(\sigma_d)$. Then, if we define (U^d,P^d) by

$$P_E{}^d = \chi_E f,$$
$$(U_z{}^d f)(x) = f(z^{-1}x),$$

(U^d,P^d) is the unique irreducible system living on E_d.

When $d=0$, $E_0=\{0\}$ and the stability group is G itself. In this case, P is trivial and U can be an arbitrary representation of G. If (U,P) is to be irreducible, U will have to be a character of G.

The example discussed just now can be generalized. Let us assume that $X=R^n$ and $G=SO(n)$, the group of orientation preserving rotations. The orbits in X are the spheres S_d of radius $d>0$ and the origin 0. Let us choose $d>0$ and consider the point

$$x_d = (0, 0, \cdots, 0, d).$$

Then the stability subgroup x_d is canonically isomorphic to $SO(n-1)$. Since G is compact, we know from corollary 6.16 that there are no ergodic systems other than the transitive ones. For each $d>0$, there are denumerably many inequivalent irreducible transitive systems living on S_d; these correspond one-one to the equivalence classes of irreducible representations of $SO(n-1)$.

However, when we proceed to describe explicitly the irreducible systems living on S_1, for example, we run into a difficulty. The theory developed in Section 5 of Chapter V and in the present chapter directs us to start with an irreducible representation, say m, of $SO(n-1)$ and construct a strict (G,S_1)-cocycle which defines the representation m at the point $(0,\cdots, 0, 1)$. Such cocycles can be constructed, of course, but they depend on the choice of a section for the left coset space $SO(n)/SO(n-1)$. Since canonical sections do not exist, our theory does not yield any simple geometric description of the unique irreducible system of imprimitivity living on S_1 which corresponds to m.

This difficulty, although of no theoretical importance, can be quite irritating when we want to make any calculation. We shall next proceed to discuss a geometric method of writing down systems of imprimitivity which does not rely on the construction of sections. This method is of course not very general, but it covers most of the cases which occur in physics.

Special Descriptions. Let X be a standard transitive Borel G-space, $x_0 \in X$. Let G_0 be the stability group at x_0. We assume that X possesses an invariant σ-finite measure, say α. Let m be any representation of G_0 in a separable Hilbert space $\mathscr{K}$. *We assume that there exists a separable Hilbert* *space* $\mathscr{K}'$ *and a Borel homomorphism* m' *of* G *into the group of invertible operators of* $\mathscr{K}'$ *such that*

$$(77)\qquad \begin{aligned} &\text{(a)}\quad \mathscr{K} \text{ is a closed linear manifold of } \mathscr{K}', \\ &\text{(b)}\quad m'(h)u = m(h)u \quad \text{for all } h \in G_0 \quad \text{and} \quad u \in \mathscr{K}. \end{aligned}$$

Let $|\cdot|$ and $\langle .,. \rangle$ denote the norm and inner product in $\mathscr{K}'$. Note that the operators $m'(g)$ for $g \in G_0$ need not be unitary. Let us consider now any point $x \in X$. Since X is transitive, there exists a $g \in G$ such that $g \cdot x_0 = x$. We now define the closed linear manifold $\mathscr{K}_x \subseteq \mathscr{K}'$ by:

$$(78)\qquad \mathscr{K}_x = m'(g)[\mathscr{K}].$$

Equation (78) defines $\mathscr{K}_x$ without ambiguity. We now introduce an inner product $\langle .,. \rangle_x$ in $\mathscr{K}_x$, by choosing $g \in G$ such that $g \cdot x_0 = x$ and putting

$$(79)\qquad \langle u,v \rangle_x = \langle m'(g^{-1})u, m'(g^{-1})v \rangle \qquad (u,\, v \in \mathscr{K}_x).$$

Once again, (79) is well defined. The fact that $\mathscr{K}_x$ and $\langle .,. \rangle_x$ are well defined and that m' is a homomorphism of G implies that

$$(80)\qquad \mathscr{K}_{g \cdot x} = m'(g)[\mathscr{K}_x],$$

$$(81)\qquad \langle m'(g)u, m'(g)v \rangle_{g \cdot x} = \langle u,v \rangle_x$$

for all $(g,x) \in G \times X$. In particular, $m'(g)$ is a *unitary isomorphism* of $\mathscr{K}_x$ onto $\mathscr{K}_{g \cdot x}$.

We now introduce the vector space $\mathscr{V}$ of α-equivalence classes of Borel maps f:

$$(82)\qquad f : X \to \mathscr{K}',$$

such that

$$(83)\qquad f(x) \in \mathscr{K}_x$$

for all $x \in X$. If $f, f' \in \mathscr{V}$ (as usual we make no distinction between functions satisfying (83) and the equivalence classes which they define), we observe that $x \to \langle f(x), f'(x) \rangle_x$ is a Borel function. To prove this, we choose a section c, i.e., a Borel map c of X into G,

$$(84)\qquad c : x \to c(x)$$

such that $c(x_0) = e$ and

$$(85)\qquad c(x) \cdot x_0 = x$$

for all x. Then,

$$(86)\qquad \langle f(x), f'(x) \rangle_x = \langle m'(c(x)^{-1})f(x), m'(c(x)^{-1})f'(x) \rangle$$

and, as the map $g \to m'(g)$ is Borel from G into the Borel space of operators on $\mathscr{K}'$, the Borelian nature of $x \to \langle f(x), f'(x) \rangle_x$ follows from corollary 6.3.

For $f \in \mathscr{V}$, we put

$$\|f\|^2 = \int_X \langle f(x),f(x)\rangle_x d\alpha(x). \tag{87}$$

Let

$$\mathscr{H} = \{f : \|f\|^2 < \infty\}. \tag{88}$$

For $f, f' \in \mathscr{H}$,

$$\langle f : f'\rangle = \int_X \langle f(x),f'(x)\rangle_x d\alpha(x) \tag{89}$$

is clearly finite. $\langle . : . \rangle$ is an inner product for $\mathscr{H}$. We define for $g \in G$ and $E \subseteq X$,

$$\begin{aligned} P_E f &= \chi_E f, \\ (U_g f)(x) &= m'(g) f(g^{-1} \cdot x), \end{aligned} \tag{90}$$

for all $f \in \mathscr{V}$. It is clear that $P_E f$ and $U_g f \in \mathscr{V}$, that P_E and U_g are linear transformations of $\mathscr{V}$, and that both U_g and P_E leave $\mathscr{H}$ invariant; it is also clear that $U_g P_E = P_{g \cdot E} U_g$ for all $g \in G$, and Borel sets $E \subseteq X$.

Theorem 6.18. *$\mathscr{H}$ is a separable Hilbert space, and (U,P), as defined by* (90), *is a system of imprimitivity in $\mathscr{H}$. Moreover, the equivalence class of (U,P) is the one induced by the representation m of G_0.*

Proof. We shall use the section $c : x \to c(x)$ satisfying (85) for all $x \in X$. With the help of c we map $\mathscr{H}$ on the Hilbert space $\mathscr{L}^2(X,\mathscr{K},\alpha)$. We define for $f \in \mathscr{V}$, Jf by

$$(Jf)(x) = m'(c(x)^{-1}) f(x). \tag{91}$$

Since $x \to f(x)$, $g \to m'(g)$, and $x \to c(x)^{-1}$ are Borel maps with appropriate domains and ranges, it follows that Jf is a Borel map of X into $\mathscr{K}$. Moreover, (86) implies that

$$\langle f(x),f(x)\rangle_x = \langle (Jf)(x),(Jf)(x)\rangle. \tag{92}$$

J is also linear and one-one. Therefore we conclude that J is a linear isomorphism of $\mathscr{H}$ onto $\mathscr{L}^2(X,\mathscr{K},\alpha)$ which preserves the inner products, i.e.,

$$\langle f : f\rangle = \langle Jf : Jf\rangle \qquad (f \in \mathscr{H}).$$

This proves that $\mathscr{H}$ is a separable Hilbert space and that J is a unitary isomorphism of $\mathscr{H}$ onto $\mathscr{L}^2(X,\mathscr{K},\alpha)$.

We then put

$$\begin{aligned} V_g &= J U_g J^{-1}, \\ Q_E &= J P_E J^{-1}. \end{aligned} \tag{93}$$

(V,Q) is a system of imprimitivity based on X, acting in $\mathscr{L}^2(X,\mathscr{K},\alpha)$. Using (90) and (91), an easy computation reveals that for $f_0 \in \mathscr{L}^2(X,\mathscr{K},\alpha)$,

$$Q_E f_0 = \chi_E f_0 \tag{94}$$

and

$$(V_g f_0)(x) = m'(c(x)^{-1}gc(g^{-1}\cdot x))f_0(g^{-1}\cdot x). \tag{95}$$

Now, the element $c(g\cdot x)^{-1}gc(x)$ lies in G_0 because it sends x_0 to $c(g\cdot x)^{-1}\cdot(g\cdot x)=x_0$. Therefore, if we define

$$\varphi(g,x) = m(c(g\cdot x)^{-1}gc(x)), \tag{96}$$

then $g, x \to \varphi(g,x)$ is a Borel map of $G \times X$ into the unitary group M of $\mathscr{K}$. Moreover, the fact that m' is a *homomorphism* leads easily to the assertion that φ is a strict (G,X,M)-cocycle. If $x=x_0$ and $g \in G_0$, then

$$\varphi(g,x) = m'(g) = m(g). \tag{97}$$

Therefore φ defines the representation m at x_0. A comparison of (95) and (96) shows that

$$(V_g f_0)(x) = \varphi(g,g^{-1}\cdot x)f_0(g^{-1}\cdot x). \tag{98}$$

Equations (97) and (98) show that (V,Q) is the system associated with m. Since $(U,P)\simeq(V,Q)$, the proof of the theorem is complete.

We shall discuss a number of situations where the conditions (77) connecting G_0, G, and m are satisfied for *every irreducible representation* m *of* G_0.

(i) *G is compact.* In this case, let m be an irreducible representation of G_0. If we decompose the representation of G, induced by m, into irreducible components, and select any irreducible, m' say, which occurs in the decomposition, it follows quite readily from the Frobenius reciprocity theorem (cf. Weil [1]) that the restriction of m' to G_0 contains m as a subrepresentation.

(ii) *G is a connected, simply connected complex semisimple Lie group, G_0 compact.* Here, G is known to be unimodular so that X admits an invariant Borel measure. From the general theory of the semisimple groups (cf. Helgason [1]), we know that there exists a *maximal* compact subgroup K containing G_0. Let m be an irreducible (finite dimensional) representation of G_0 acting in $\mathscr{K}$. By (i) above, there exists a finite dimensional Hilbert space $\mathscr{K}'$ containing $\mathscr{K}$ and an irreducible representation m' of K in $\mathscr{K}'$ such that $m'(h)=m(h)$ for all $h \in G_0$. Now, it is well known (cf. Weyl [2], p. 267) that there exists a homomorphism m'' of G into the group of invertible operators of $\mathscr{K}'$ *itself* such that $m''(k)=m'(k)$ for $k \in K$, and that m'' is even *unique* if we require that the map $g \to m''(g)$ is *complex analytic on* G. Thus the conditions of theorem 6.18 are satisfied in this case. In particular, if G_0 *itself is maximal compact in* G, the systems of imprimitivity based on G/G_0 can be described in the geometric form of theorem 6.18. We shall do this in Chapter IX, when $G = SL(2,\mathbf{C})$ and $G_0 = SU(2,\mathbf{C})$, in connection with the equations for the Dirac electron. We thus have:

Theorem 6.19. *Let* G *be a connected, simply connected complex semisimple Lie group, and* G_0 *a maximal compact subgroup of* G. *Let* α *be an invariant Borel measure on* $X = G/G_0$. *Let* m *be an irreducible representation of* G_0 *in a Hilbert space* $\mathscr{K}$, *and let* $m'(g \to m'(g))$ *be the unique complex-analytic homomorphism of* G *into the group of invertible linear endomorphisms of* $\mathscr{K}$ *such that* $m'(h) = m(h)$ *for all* $h \in G_0$. *Then there exists exactly one map* $x \to \langle \cdot,\cdot \rangle_x$ *of* X *into the set of positive definite inner products on* $\mathscr{K} \times \mathscr{K}$ *such that* $\langle \cdot,\cdot \rangle_{x_0} = \langle \cdot,\cdot \rangle$ *and*

$$\langle u,v \rangle_x = \langle m'(g)u, m'(g)v \rangle_{g \cdot x}$$

for all $(g,x) \in G \times X$. *If* $\mathscr{H}$ *is the Hilbert space of Borel maps* f *of* X *into* $\mathscr{K}$ *such that*

$$\|f\|^2 = \int_X \langle f(x), f(x) \rangle_x d\alpha(x) < \infty, \tag{99}$$

and we define P *and* U *by*

$$\begin{aligned} P_E f &= \chi_E f, \\ (U_g f)(x) &= m'(g) f(g^{-1} \cdot x), \end{aligned} \tag{100}$$

then (U,P) *is a system of imprimitivity equivalent to the system induced by* m.

Representations in Vector Bundles. The reader might have noticed that the essential point of the proof of theorem 6.18 is contained in the fact that the family of subspaces $x \to \mathscr{K}_x$ transform "covariantly" under the action of G. It is natural to suspect that a more general formulation of theorem 6.18 can be obtained in the context of such covariant families. This is indeed so, and the natural setup for formulating this is that of a *vector bundle*, more precisely, a *Hilbert space bundle*.

Let X and B be both standard Borel G-spaces and let X be transitive. We assume that there is a Borel map

$$\pi : B \to X \tag{101}$$

of B onto X with the property that for each $g \in G$, the diagram

$$\begin{tikzcd} B \arrow[r, "g"] \arrow[d, "\pi"'] & B \arrow[d, "\pi"] \\ X \arrow[r, "g"'] & X \end{tikzcd} \tag{102}$$

is commutative. If we write, for each $g \in G$, $D(g)$ for the automorphism of B induced by g, and $x \to g \cdot x$ for the automorphism of X induced by g, (102) translates to

$$\pi(D(g)(b)) = g \cdot \pi(b) \tag{103}$$

for each $b \in B$ and each $g \in G$. We shall assume also for simplicity that X admits an invariant σ-finite measure, say α.

We shall say that B is a *G-Hilbert space bundle* if, for each $x \in X$,

$$B_x = \pi^{-1}(\{x\}) \tag{104}$$

is a separable Hilbert space whose natural Borel structure is the one induced on B_x by B and if, for each $g \in G$ and $x \in X$,

$$b \to D(g)b \qquad (b \in B_x) \tag{105}$$

is a *unitary* isomorphism of B_x onto $B_{g \cdot x}$. If we write $|\cdot|_x$ and $\langle .,. \rangle_x$ for the norm and inner product on B_x, (105) means

$$\langle b,b' \rangle_x = \langle D(g)b, D(g)b' \rangle_{g \cdot x} \qquad (b, b' \in B_x). \tag{106}$$

In particular, if G_x is the stability subgroup at x, $D(g)$ leaves B_x invariant for $g \in G_x$ and $g \to D(g)$ defines a representation of G_x in B_x.

We shall now consider sections of the bundle. A *section* of B is a Borel map f,

$$f : x \to f(x) \tag{107}$$

of X into B such that $f(x) \in B_x$ for all $x \in X$. To handle the measurability problems involving these sections, we transform them into Borel maps of X into a single Hilbert space. Choose a point $x_0 \in X$ and a Borel map c from X into G such that $c(x_0) = e$ and

$$c(x) \cdot x_0 = x$$

for all $x \in X$. Since the map $g, b \to D(g)b$ of $G \times B$ into B is Borel, it follows quite simply that for any section f of B, $x \to D(c(x))^{-1}f(x)$ is a Borel map of X into B_{x_0}. Conversely, if $x \to f_0(x)$ is a Borel map of X into B_{x_0}, then $x \to D(c(x))f_0(x)$ is a section of the bundle B. We denote by $\mathscr{V}$ the vector space of α-equivalence classes of sections.

For any $x \in X$, and any section f,

$$\langle f(x), f(x) \rangle_x = \langle D(c(x))^{-1}f(x), D(c(x))^{-1}f(x) \rangle_{x_0}.$$

This shows that $x \to |f(x)|_x^2$ is a Borel function on X. We shall say that f is a square integrable section if

$$\|f\|^2 = \int_X |f(x)|_x^2 d\alpha(x) < \infty. \tag{108}$$

If $\mathscr{H}$ denotes the subset of $\mathscr{V}$ consisting of square integrable sections, then $\mathscr{H}$ is a pre-Hilbert space under the inner product

$$\langle f : f' \rangle = \int_X \langle f(x), f'(x) \rangle_x d\alpha(x). \tag{109}$$

The isomorphism, which associates with each section f, the map $x \to D(c(x))^{-1} f(x)$, induces a unitary isomorphism W of $\mathscr{H}$ onto $\mathscr{L}^2(X, B_{x_0}, \alpha)$. $\mathscr{H}$ is thus a separable Hilbert space.

For any Borel set $E \subseteq X$ we write

$$P_E f = \chi_E f \qquad (f \in \mathscr{H}). \tag{110}$$

For each $g \in G$, we define the operator U_g by

$$(U_g f)(x) = D(g) f(g^{-1} \cdot x) \qquad (f \in \mathscr{H}). \tag{111}$$

Computations similar to those of theorem 6.18 then give rise to the following theorem.

Theorem 6.20. *(U, P) is a system of imprimitivity. If D^0 is the representation of the stability subgroup G_{x_0} (of G at x_0) in the Hilbert space B_{x_0}, then (U, P) is equivalent to the system induced by D^0.*

As an interesting example of representations in such bundles we mention the following. Let X be a C^∞ Riemannian manifold, and let each $g \in G$ be an *isometry* of X. We assume that G is a Lie group acting transitively on X in such a fashion that $g, x \to g \cdot x$ is a C^∞-map of $G \times X$ into X. Let B be the *tangent bundle*† *of* X. For each g, let $D(g)$ be the differential of the isometry $x \to g \cdot x$. Then all the conditions of the preceding theorem are satisfied. The system of imprimitivity thus acts on the Hilbert space of square integrable vector fields and is equivalent to the system induced by the representation of the stability group G_0 at $x_0 \in X$ defined in the tangent space† to X at x_0.

In Chapter IX, we shall see how the representations associated with the *photon* can be formulated, in the spirit of theorem 6.20, as acting in the Hilbert space of sections of a certain vector bundle on the light cone, leading thereby in a very simple way to the equations of Maxwell.

Ergodic Intransitive Systems. The last example is simply that of a standard Borel G-space which has ergodic nontransitive (invariant) measure classes. We take real numbers $\lambda_1, \lambda_2, \cdots$ which are rationally independent and take X to be a torus having its dimension equal to the number of λ's (finite or infinite). Let $G = \mathbf{R}^1$, and for $t \in \mathbf{R}^1$ and

$$x = (\zeta_1, \zeta_2, \cdots) \qquad (|\zeta_j| = 1 \text{ for all } j) \tag{112}$$

---

† Complexified.

in X, let

$$t \cdot x = (e^{it\lambda_1}\zeta_1, e^{it\lambda_2}\zeta_2, \cdots). \tag{113}$$

Then, it is classical (Kronecker) that each orbit is dense and that the Haar measure on X is invariant and ergodic but not transitive.

8. SEMIDIRECT PRODUCTS

We shall now use the theory of systems of imprimitivity to obtain the description of *all* irreducible representations of an interesting class of lcsc groups which are neither compact nor abelian. This class of groups includes the inhomogeneous Lorentz group and the group of Euclidean motions as members. For the special case of the Lorentz group, the analysis which we shall give now was carried out by Wigner in his 1939 paper. The general study is due to Mackey and comes out as a simple consequence of the work in Section 6.

We begin by introducing the concept of semidirect products. Let A and H be two groups and for each $h \in H$ let

$$t_h : a \to h[a] \tag{114}$$

be an automorphism of the group A. We shall assume that $h \to t_h$ is a homomorphism of H into the group of automorphisms of A so that

$$h[a] = a \text{ for all } a \in A \text{ if } h = e_H, \text{ the identity of } H,$$
$$t_{h_1h_2} = t_{h_1}t_{h_2} \qquad (h_1,h_2 \in H). \tag{115}$$

Equation (115) converts A into an H-space. We shall now make

$$G = H \times A$$

into a *group* by defining the multiplication in G by

$$(h,a)(h',a') = (hh',at_h[a']). \tag{116}$$

It is easy to verify that this definition converts G into a group with $e=(e_H,e_A)$ as its identity (e_A being the identity of A). Further,

$$(h,a)^{-1} = (h^{-1},h^{-1}[a^{-1}]). \tag{117}$$

G is called the *semidirect product of H and A relative to t* and we put

$$G = H \times_t A. \tag{118}$$

When no confusion can arise as to what t is, we omit it, and write $G=H\times' A$. But it must be remembered that (i) A and H play very unlike roles in the formation of G, in contrast to the construction of direct products, and (ii) G can be formed only after specifying t, i.e., specifying how H "acts" on A.

Suppose now that A and H are lcsc groups and that the map

$$h, a \rightarrow h[a] \tag{119}$$

is continuous from $H \times A$ into A. Then, if we equip G with the product topology, it follows quite easily from (116) and (117) that G becomes a topological group. Clearly, G is a lcsc group also. We shall denote this topological group also by $H \times_t A$. We shall fix A, H, and t throughout this section.

A quick calculation shows that

$$(h,a)(h',a')(h,a)^{-1} = (hh'h^{-1},ah[a']hh'h^{-1}[a^{-1}]). \tag{120}$$

It follows from (120) that

$$A^{\sim} = \{(e_H,a')\} \tag{121}$$

is a closed normal subgroup of G and that

$$(h,a)(e_H,a')(h,a)^{-1} = (e_H,ah[a']a^{-1}). \tag{122}$$

In particular, the inner automorphism of G induced by (h,e_A) coincides on A with the automorphism $(e_H,a') \rightarrow (e_H,h[a'])$. We put

$$H^{\sim} = \{(h,e_A)\}; \tag{123}$$

then $H^{\sim}$ is a closed subgroup of G. Identifying A with $A^{\sim}$ and H with $H^{\sim}$ we find

$$\begin{aligned} G &= AH, \\ \{e\} &= A \cap H, \\ h[a] &= hah^{-1}. \end{aligned} \tag{124}$$

The inhomogeneous linear groups are typical examples of semidirect products. In these examples, A is a real (or complex) finite dimensional vector space and $a \rightarrow h[a]$ is an invertible linear endomorphism of A. H is thus identified with a subgroup of the group of all invertible (linear) endomorphisms of A, which is usually closed. The Euclidean group and the inhomogeneous Lorentz group fit into this general framework. For the Euclidean group, $A=\mathbf{R}^n$ and H is the group $SO(n)$. For the inhomogeneous Lorentz group, $A=\mathbf{R}^4$ and H is the group of all invertible linear transformations of $\mathbf{R}^4$ which preserve the quadratic form $x_0{}^2-x_1{}^2-x_2{}^2-x_3{}^2$ (x_0, x_1, x_2, x_3 are coordinates on A). It undoubtedly would be of great interest therefore to develop a theory of the representations of semidirect products. In complete generality this has not been done. However, under certain restrictive hypotheses, we shall construct a theory of representations of semidirect products which goes far enough to yield all physically important results. *From now on, through the rest of* *the section, we shall assume that A is a closed, abelian, normal subgroup of G, H a closed subgroup of G such that the conditions* (124) *are satisfied.*

The first step of our theory is to obtain the representations of A. The irreducible representations are the characters of A, i.e., continuous homomorphisms of A into the group of complex numbers of modulus 1. Let $\hat{A}$ be the set of all characters of A. Under pointwise multiplication, $\hat{A}$ becomes a group also. We shall equip $\hat{A}$ with the topology of uniform convergence on compacta. Then it can be shown that $\hat{A}$ becomes a topological group and is in fact a locally compact Abelian group satisfying the second axiom of countability (cf. Pontrjagin [1]). Let us then take a σ-finite measure α on $\hat{A}$ and consider the Hilbert space $\mathscr{H} = \mathscr{L}^2(\hat{A},\alpha)$. For each $a \in A$ and $f \in \mathscr{H}$ let $(U_a f)(x) = x(a)f(x)(x \in \hat{A})$, where $x(a)$ is the value of the character x at a. Then $a \rightarrow U_a$ is a representation of A. These representations do not exhaust the representations of A. To obtain more general representations we replace α by a *projection valued measure* P based on $\hat{A}$. Let $\mathscr{H}$ be a (separable) Hilbert space and P a projection valued measure based on $\hat{A}$ acting in $\mathscr{H}$. For each $a \in A$ let U_a be the operator in $\mathscr{H}$ defined by

$$\langle U_a f, f' \rangle = \int_{\hat{A}} x(a) d\nu_{f,f'}(x),$$

where

$$\nu_{f,f'}(E) = \langle P_E f, f' \rangle.$$

We shall symbolically write

$$U_a = \int_{\hat{A}} x(a) dP(x). \tag{125}$$

It is easy to show that U_a is unitary and that $a \rightarrow U_a$ is a representation of A in $\mathscr{H}$. Conversely, it is a known theorem that if U is any representation of A in $\mathscr{H}$, there exists a *unique* projection valued measure P based on $\hat{A}$ such that

$$U_a = \int_{\hat{A}} x(a) dP(x) \tag{126}$$

for all $a \in A$. P is called *the projection valued measure corresponding to* U. The formula (126) shows how U can be constructed from P. Thus, for instance, it is quite clear that if T is an operator commuting with all P_E, then T commutes with all U_a. Conversely, let T commute with all U_a. Then T commutes with all the spectral projections of all the U_a. Now, if η_a is the continuous map $x \rightarrow x(a)$ of $\hat{A}$ into the group M of complex numbers of modulus 1, it follows from (126) that $Q^a(F \to P_{\eta_a^{-1}(F)})$ is a projection valued measure based on M and

$$U_a = \int_M z dQ^a(z).$$

This shows that the $P_{\eta_a^{-1}(F)}$ are the spectral projections of U_a and hence that T commutes with all of them. Since the sets of the form $\eta_a{}^{-1}(F)$ generate the Borel structure of $\hat{A}$, it follows that T commutes with P. The reader who wants to study the (Fourier) analysis of representations of A may consult Loomis [2, Chapter VII].

Let us now consider the lcsc group G and a representation U of G in a separable Hilbert space $\mathscr{H}$. Our method of analysis of U is essentially the Fourier analysis of the restriction of U to A. We begin this analysis by first constructing the "adjoint" action of H on $\hat{A}$.

Lemma 6.21. *Let $h \in H$. Then, for each $x \in \hat{A}$ there exists one and only one $y \in \hat{A}$ such that*

$$y(a) = x(h^{-1}[a]) \tag{127}$$

for all $a \in A$. If we write $y = h[x]$, then $h, x \to h[x]$ is continuous from $H \times \hat{A}$ into $\hat{A}$ and $\hat{A}$ becomes a H-space.

Proof. That y is unique and well defined is obvious. The remaining assertions are easy to prove. We omit the proofs.

Not every representation of A in $\mathscr{H}$ can be enlarged to a representation of G in $\mathscr{H}$. Our next question is: when can this be done?

Lemma 6.22. *Let U and V be representations of A and H, respectively, in a separable Hilbert space $\mathscr{H}$ and let P be the projection valued measure on $\hat{A}$ corresponding to U. Then, a necessary and sufficient condition that there should exist a representation W of G in $\mathscr{H}$ whose restrictions to A and H are U and V, respectively, is that (V,P) is a system of imprimitivity for H, based on $\hat{A}$. In this case, W is unique.*

Proof. Let W be a representation of G in $\mathscr{H}$ and let $U_a = W_a$ $(a \in A)$, $V_h = W_h$ $(h \in H)$. U and V are representations of A and H, respectively. Now

$$hah^{-1} = h[a]$$

so that

$$V_h U_a V_h{}^{-1} = U_{h[a]} \tag{128}$$

for all $(h,a) \in H \times A$. Let P be the projection valued measure on $\hat{A}$ corresponding to U. Now, a routine calculation based on transport of structure shows that the projection valued measure of the representation $a \to V_h U_a V_h{}^{-1}$ of A is $E \to V_h P_E V_h{}^{-1}$, and that corresponding to the representation $a \to U_{h[a]}$ is $E \to P_{h[E]}$. In view of the uniqueness of the projection valued measures which correspond to representations of A we infer from (128) that

$$V_h P_E V_h^{-1} = P_{h[E]}. \tag{129}$$

(V,P) is thus a system of imprimitivity for H based on $\hat{A}$.

Conversely, let us start with U, V, and P such that P and V satisfy (129). Then U and V satisfy (128). Define W on G by

$$W_{ah} = U_a V_h. \tag{130}$$

The relations (128) are now enough to secure the fact that W is a homomorphism. Since $a \to U_a$ and $h \to V_h$ are Borel, $x \to W_x$ is Borel. W is thus a representation of G and its restrictions to H and A are V and U, respectively. Since $G = AH$, W is unique. The lemma is proved.

The lemma just proved enables us to establish a one-one correspondence between representations of G and systems of imprimitivity of H based on $\hat{A}$.

Lemma 6.23. *A representation W of G is irreducible if and only if the corresponding system of imprimitivity for H based on $\hat{A}$ is irreducible. Two representations of G are equivalent if and only if the corresponding systems of imprimitivity are equivalent.*

Proof. Let W be a representation of G in $\mathscr{H}$ and let (V,P) be the corresponding system of imprimitivity, also acting in $\mathscr{H}$. We use the notation of lemma 6.22. If T is an operator in $\mathscr{H}$, then T commutes with all the U_a $(a \in A)$ if and only if T commutes with all P_E. From this it follows that T lies in the commuting ring of W if and only if T lies in the commuting ring of (V,P). This proves the first assertion. For the second, let W^i be a representation of G in $\mathscr{H}^i$, and (V^i,P^i) the corresponding system of imprimitivity $(i=1, 2)$. If T is a unitary isomorphism of $\mathscr{H}^1$ onto $\mathscr{H}^2$, it follows from the uniqueness of the correspondence $U^i \rightleftarrows P^i$ that $U_a{}^2 = TU_a{}^1T^{-1}$ for all $a \in A$ if and only if $P_E{}^2 = TP_E{}^1T^{-1}$ for all Borel sets $E \subseteq \hat{A}$. From this the second assertion of the lemma follows quickly.

The preceding lemmas and the theory of Sections 5 and 6 reduce the problem of describing the irreducible representations of G to the problem of describing the irreducible representations of the various stability subgroups of H with respect to its action on $\hat{A}$. In view of the applications to the study of relativistic equations we proceed to spell out the details of the relevant constructions.

We choose a point $x_0 \in \hat{A}$ and a σ-finite measure α which is quasi invariant on the H-space $\hat{A}$ and which lives on the orbit $H[x_0] = X$. Let $H_{x_0} = H_0$ be the stability group at x_0, i.e.,

$$H_0 = \{h : h \in H, \quad h[x_0] = x_0\}, \tag{131}$$

and let m be an irreducible representation of H_0 in a separable Hilbert space $\mathscr{K}$. Let us define $\mathscr{H} = \mathscr{L}^2(\hat{A},\mathscr{K},\alpha)$ and let (V,P) be a system of imprimitivity for H based on $\hat{A}$, living on the orbit X, and induced by the representation m. We put, for $ah \in G$ $(a \in A,\, h \in H)$ and $f \in \mathscr{H}$,

$$(132) \qquad (W_{ah}f)(x) = x(a)(V_h f)(x) \qquad (x \in \hat{A}).$$

Since $P_E f = \chi_E f$ for all Borel sets $E \subseteq \hat{A}$, it follows easily that, for the representation $a \to U_a$ of A corresponding to P,

$$(133) \qquad (U_a f)(x) = x(a) f(x) \qquad (x \in \hat{A}),$$

so that $W_{ah} = U_a V_h$. Thus W is a representation of G. From our work in Sections 5 and 6 we know that the equivalence class of the system (V,P) depends only on the measure class of α and the equivalence class of m. From lemmas 6.22 and 6.23 it follows that W is irreducible. We shall say that W is *associated* with x_0 and m. The following theorem is now an immediate consequence of our work in Sections 5 and 6 of the present chapter.

Theorem 6.24 (*Mackey* [2] [3]). *Let us choose, for each H-orbit* ω *in* $\hat{A}$, *one point* x_ω *on* ω, *and an irreducible representation* m, *of the stability subgroup of* H_ω *at the point* x_ω. *Then the representation* $W^{m,\omega}$ *of G associated with* x_ω *and* m *is irreducible.* $W^{m,\omega}$ *and* $W^{m',\omega'}$ *are equivalent if and only if the orbits* ω *and* ω' *coincide, and the representations* m *and* m' *are equivalent. If the H-orbit structure of* $\hat{A}$ *is smooth, then each irreducible representation of G is equivalent to some* $W^{m,\omega}$.

NOTES ON CHAPTER VI

The theory of induced representations is originally due to Frobenius who studied it for finite groups. Its generalization to the category of locally compact second countable groups is due to Mackey. Nowadays Mackey's theory is referred to as the *Mackey Machine*; it is also known as the *little group method* among physicists. Significant contributions were also made by I. M. Gel'fand and M. A. Naĭmark in their work on representations of complex classical groups (cf. Trudy Mat. Inst. Steklova, **36** (1950), pp. 1–288), and by F. Bruhat (Bull. Soc. Math. France, **84** (1956), pp. 97–205) who developed the theory of induced representations for Lie groups with the techniques of the theory of Schwartz distributions. Especially noteworthy are the various expositions of Mackey in: *The Theory of unitary group representations in Physics, Probability, and Number Theory*, Benjamin/Cummings, Reading, Mass., 1978; *Harmonic analysis as the exploitation of symmetry—a historical survey*, Bull. Amer. Math. Soc., **3** (1980), pp. 543–698.

Unitary representation theory is so vast and currently so active that it is not possible to begin to sketch its outlines. For some notable surveys see: I. M. Gel'fand, Proceedings of the International Congress of Mathematicians Stockholm, 1962, pp. 74–85; the articles in *Harmonic Analysis on homogeneous spaces*, Proceedings of Symposia in Pure Mathematics, Vol. XXVI, Amer. Math. Soc. Providence, R.I., (ed.) C.C. Moore, 1973; and V.S. Varadarajan, *Harmonic analysis on real semisimple Lie groups*, Proceedings of the International Congress of Mathematicians, Vancouver, 1974, pp. 121–127. The theory for semisimple groups is mainly the creation of Harish-Chandra; for this, see Harish-Chandra, *Collected Papers*, 4 Vols., Springer-Verlag, New York, 1984; see also *Harmonic analysis and representations of semisimple Lie groups*, Reidel, Dordrecht, 1980, edited by J. A. Wolf, M. Cahen, and M. de Wilde. Recently, interest has intensified in representation theory of infinite dimensional groups; for a view through the theory of infinite dimensional Lie algebras see the book of Kac and the references therein: V. G. Kac, *Infinite dimensional Lie algebras*, Birkhäuser, Boston, 1983.

CHAPTER VII

MULTIPLIERS

1. THE PROJECTIVE GROUP

Earlier in Chapter III we showed that in any mathematical description of a quantum mechanical system the requirement of covariance introduces a representation of the symmetry group G of the system into the group Aut($\mathscr{S}$), where $\mathscr{S}$ is the set of all states of the system in question. If we assume that the logic of the system is standard, this representation can be replaced by a representation of G into the group of all symmetries of the Hilbert space underlying the logic. If G^0 is the subgroup of G consisting of all elements in the same connected component as the identity of G, then it can be shown that when G is a Lie group the symmetry corresponding to each element of G^0 is a unitary rather than an antiunitary operator†. However, these unitary operators are not uniquely determined. Each of them can obviously be multiplied by a complex number of modulus 1 (called a phase factor) without changing the induced automorphism of the logic. It is therefore not immediately obvious that we have a *unitary representation* of G^0, i.e., that the phase factors involved are all removable.

In this chapter we shall examine this question. We remark that it was first studied by Wigner [1]. He proved, among other things that, if G is the inhomogeneous Lorentz group, *any* representation of G in Aut($\mathscr{S}$) can be induced by a unitary representation, at least of the *universal covering group* of G. This result was the starting point for his classification of relativistic wave equations. Wigner's discussion of this problem, however, was restricted to the Lorentz and rotation groups. It was Bargmann [1] who first examined these problems systematically. He obtained many general theorems which included Wigner's results as special cases.

In his work on group representations, Mackey was led to some of these problems from a purely mathematical point of view. In his paper [6] he formulated the basic concepts of a theory of such phase factors for arbitrary lcsc groups and showed how the subject can be fitted into the general framework of the theory of group extensions. This is the point of view taken in this chapter. With every phase factor, or multiplier as we call them, we shall, following Mackey, associate a group extension of the unit circle by the given group and reduce the solution of problems involving that multiplier to problems involving this group extension. The subject then becomes essentially a part of a cohomology theory of certain types of extensions.

---

† If $x \in G$ is of the form y^2 for $y \in G$, then the symmetry associated with x is necessarily unitary. Now, if G is a Lie group and G^0 the connected component of the identity, then for any x of the form $\exp X$ (X in the Lie algebra of G), $x = y^2$, where $y = \exp \frac{1}{2}X$. Since the range of the map $X \rightarrow \exp X$ contains a neighborhood of the identity, it follows that every element in a certain neighborhood N of the identity of G is mapped into a unitary operator in a given representation. Thus the elements of the group generated by the elements of N are also mapped into unitary operators. But G^0 coincides with this group.

Our purpose is not to develop a general theory of group extensions. Such a general discussion may be found, in various contexts and at widely differing levels, in the works of Hochschild [1], [2], Calabi [1], Moore [1], Mackey [7], [8], Kleppner [1], and a number of other authors. Our aim is the much more modest one of developing in some detail the basic concepts of group extensions, so as to be able to obtain the solutions to the problems arising in connection with the physical space-time symmetry groups.

We might mention that Hermann Weyl (cf. [1]), has also emphasized the fact that it is the automorphism of the logic and not the underlying symmetry that is of significance in physics. Some of the facts about projective representations may be found in his book.

Let $\mathscr{H}$ be a complex separable Hilbert space. We denote by $\mathscr{U}$ the unitary group of $\mathscr{H}$. If we equip $\mathscr{U}$ with the strong topology, we have seen in Chapter V that $\mathscr{U}$ becomes a metrizable second countable topological group and that its Borel structure is standard. We denote by $\mathscr{Z}$ the set of all elements of $\mathscr{U}$ which are multiples of the identity operator. $\mathscr{Z}$ is a closed, normal, subgroup of $\mathscr{U}$ and is *central*, i.e., each element of $\mathscr{Z}$ commutes with all the elements of $\mathscr{U}$. The group

$$\mathscr{P} = \mathscr{U}/\mathscr{Z} \tag{1}$$

is called the *projective group of* $\mathscr{H}$. We write π for the canonical homomorphism

$$\pi : U \to \pi(U) \qquad (U \in \mathscr{U}) \tag{2}$$

of $\mathscr{U}$ onto $\mathscr{P}$. Theorem 4.27 now tells us that the subgroup of the group of automorphisms of the logic of $\mathscr{H}$, consisting of automorphisms which can be induced by unitary operators, is isomorphic to $\mathscr{P}$.

We shall equip $\mathscr{P}$ with the quotient topology. Thus, a set $A \subseteq \mathscr{P}$ is defined to be open if and only if $\pi^{-1}(A)$ is an open subset of $\mathscr{U}$. It is easy to show that $\mathscr{P}$ becomes a metrizable topological group satisfying the second axiom of countability and that π is an open continuous map of $\mathscr{U}$ onto $\mathscr{P}$. If we write

$$T = \{z : z \in \mathbf{C}, \quad |z| = 1\}, \tag{3}$$

then the kernel of π is precisely the set of operators of the form $z1$ ($z \in T$), and is hence a compact group. In all statements involving the topology and Borel structure on $\mathscr{P}$, we shall refer only to the above-described topology of $\mathscr{P}$ and its natural Borel structure.

Lemma 7.1. *The following statements concerning a sequence* $\{U_n\}_{n=0,1,2,\cdots}$ *of elements in* $\mathscr{U}$ *are equivalent:*

(i) $|\langle U_n f, f'\rangle| \to |\langle U_0 f, f'\rangle|$ *for all* $f, f' \in \mathscr{H}$,

(ii) *there exists a sequence* $\{z_n\}$ *of complex numbers of modulus* 1 *such that* $z_n U_n \to U_0$ *in* $\mathscr{U}$,

(iii) $\pi(U_n) \to \pi(U_0)$ *in* $\mathscr{P}$.

Proof. (ii) $\Rightarrow$ (iii) obviously. On the other hand, for fixed $f, f' \in \mathscr{H}$, $U \to |\langle Uf, f'\rangle|$ is a continuous function on $\mathscr{U}$ which is constant on each $\mathscr{Z}$-coset. Hence, by virtue of the fact that $\mathscr{P}$ has the quotient topology, it defines a continuous function on $\mathscr{P}$. The implication (iii) $\Rightarrow$ (i) follows from the assertion that this function is continuous on $\mathscr{P}$. We shall complete the proof of the lemma by showing that (i) $\Rightarrow$ (ii).

We begin the proof with a simple remark. Suppose U is an operator of $\mathscr{H}$ such that $|\langle Uf, f'\rangle| = |\langle f, f'\rangle|$ for all $f, f' \in \mathscr{H}$. Then $U = t1$ for some $t \in T$. In fact, for any f, $Uf \in \{f\}^{\perp\perp}$, so that Uf is a multiple of f. Since this is true for each f, U must be a multiple of the identity, $U = t1$. It is obvious that $|t| = 1$, and thus $t \in T$.

This said, we proceed to show that (i) $\Rightarrow$ (ii). Replacing U_n by $U_n U_0^{-1}$ we may assume that $U_0 = 1$. Since $\|U_n\| \leq 1$, there exists an operator U with $\|U\| \leq 1$ and a subsequence $\{U_{n_k}\}$ such that $\langle U_{n_k} f, f'\rangle \to \langle Uf, f'\rangle$ for all $f, f' \in \mathscr{H}$ (Dunford-Schwartz [1]). From our remark in the previous paragraph, we conclude that $U = t1$ for some $t \in T$. Since each U_n is unitary, $U_{n_k} f \to tf$ for each f so that $U_{n_k} \to t1$. Since we can apply this argument to any subsequence of $\{U_n\}$, it follows that $\{U_1, U_2, \cdots\}$ is a set with compact closure and that its only limit points are of the form $t1$ ($t \in T$). Choose a unit vector f and write (for sufficiently large n) $z_n = \langle U_n f, f\rangle^{-1} |\langle U_n f, f\rangle|$. We claim that 1 is the only limit point of the sequence $\{z_n U_n\}$. If $z_{n_k} U_{n_k} \to V$, then $V = t1$ for some $t \in T$, and as $\langle z_{n_k} U_{n_k} f, f\rangle$ is real and >0 for all sufficiently large k, $t = 1$. Since $\{z_1 U_1, z_2 U_2, \cdots\}$ has compact closure, we see that $z_n U_n \to 1$.

For $f, f' \in \mathscr{H}$ and $V \in \mathscr{U}$, we write $d_{f,f'}(\pi(V)) = |\langle Vf, f'\rangle|$.

Corollary 7.2. *The topology for* $\mathscr{P}$ *is the smallest one with respect to which the maps* $d_{f,f'}$ *are all continuous. If* Z *is a Borel space and* $z \to U(z)$ *is a map of* Z *into* $\mathscr{U}$, *then* $z \to \pi(U(z))$ *is a Borel map of* Z *into* $\mathscr{P}$ *if and only if* $z \to |\langle U(z) f, f'\rangle|$ *is Borel for all* $f, f' \in \mathscr{H}$.

Proof. Let $\mathscr{T}$ be the smallest topology in question. We have already seen that $|\langle Vf, f'\rangle| = |\langle Wf, f'\rangle|$ for all $f, f' \in \mathscr{H}$ if and only if $\pi(V) = \pi(W)$.

$\mathscr{T}$ is hence Hausdorff. Let D be a countable dense subset of $\mathscr{H}$. Then an easy argument shows that $\mathscr{T}$ is also the smallest topology with respect to which the maps $d_{f,f'}$ $(f, f' \in D)$ are continuous. Hence $\mathscr{T}$ is metrizable and lemma 7.1 now implies the first statement. The second statement is now clear.

Lemma 7.3. *Let $f \in \mathscr{H}$ be nonzero and let*

$$Q_f = \{U : U \in \mathscr{U}, \quad \langle Uf,f\rangle \text{ is real and } > 0\}. \tag{4}$$

Then Q_f is a Borel set in $\mathscr{U}$ containing 1 which meets each $\mathscr{Z}$-coset at most once. Moreover, $\pi[Q_f]$ is open in $\mathscr{P}$ and π is a homeomorphism of Q_f onto $\pi[Q_f]$.

Proof. That Q_f is a Borel set containing 1 is obvious. If $\langle Uf,f\rangle > 0$ and $t \in T$, $\langle tUf,f\rangle > 0$ if and only if $t = 1$. Thus Q_f meets each $\mathscr{Z}$-coset at most once. On the other hand, if $\langle Uf,f\rangle = t \neq 0$, then for $V = |t|t^{-1}U$, one has $\langle Vf,f\rangle = |t| > 0$. This proves that

$$\pi^{-1}(\pi[Q_f]) = \{U : U \in \mathscr{U}, \quad \langle Uf,f\rangle \neq 0\}. \tag{5}$$

Equation (5) shows at once that $\pi[Q_f]$ is open in $\mathscr{P}$. It remains to prove that π is a homeomorphism on Q_f. Let $\{U_n\}_{n=0,1,\cdots}$ be a sequence in Q_f such that $\pi(U_n) \to \pi(U_0)$ in $\mathscr{P}$. By lemma 7.1 there exists a sequence $\{t_n\}$ in T such that $t_n U_n \to U_0$ in $\mathscr{U}$. Since $|\langle U_n f,f\rangle| = \langle U_n f, f\rangle \to |\langle U_0 f,f\rangle| = \langle U_0 f,f\rangle$, it follows that $\{t_n\}$ itself converges to 1. This shows that $U_n \to U_0$. As π is continuous, this proves that π is a homeomorphism and completes the proof of the lemma.

Theorem 7.4. *$\mathscr{P}$ is a standard Borel group. There exists a Borel map c from $\mathscr{P}$ into $\mathscr{U}$,*

$$c : x \to c(x) \tag{6}$$

such that

$$\begin{aligned} &\text{(i)} \quad \pi(c(x)) = x \quad \textit{for all} \quad x \in \mathscr{P}, \\ &\text{(ii)} \quad c(\pi(1)) = 1, \\ &\text{(iii)} \quad \textit{there is an open set containing } \pi(1) \textit{ on which } c \textit{ is continuous.} \end{aligned} \tag{7}$$

The range of c is a Borel subset of $\mathscr{U}$ which meets each $\mathscr{Z}$-coset exactly once. Finally, if f is a map of $\mathscr{P}$ into some Borel space X, f is Borel if and only if the map $U \to f(\pi(U))$ of $\mathscr{U}$ into X is Borel. In particular, $A \subseteq \mathscr{P}$ is Borel if and only if $\pi^{-1}(A)$ is a Borel set in $\mathscr{U}$.

Proof. We begin our proof with the construction of c. Let $\{f_1, \cdots, f_n, \cdots\}$ be a countable dense set of nonzero vectors and let (cf. (4))

$$Q_n = Q_{f_n}. \tag{8}$$

We now define sets $E_1, E_2, \cdots \subseteq \mathscr{U}$ as follows:

$$E_1 = Q_1, \tag{9}$$

and for $n > 1$,

$$E_n = Q_n \cap \bigcap_{i=1}^{n-1} \{U : \langle Uf_i, f_i\rangle = 0\}. \tag{10}$$

Let

$$E = \bigcup_{n=1}^{\infty} E_n. \tag{11}$$

E is a Borel subset of $\mathscr{U}$. We claim that E meets each $\mathscr{Z}$-coset exactly once. In fact, let $V \in \mathscr{U}$ and let m be the smallest of integers i such that $\langle Vf_i, f_i\rangle \neq 0$; it is then obvious that $V\mathscr{Z} \cap E = V\mathscr{Z} \cap E_m$ consists of one element. E is thus a Borel set meeting each $\mathscr{Z}$-coset exactly once.

The canonical map π is thus one-one on E and maps E onto $\mathscr{P}$. Since $\mathscr{P}$ is a separable Borel space and E is standard, the image of any Borel subset $F \subseteq E$ under π is a Borel set in $\mathscr{P}$. Hence π is a Borel isomorphism on E. In particular, $\mathscr{P}$ is a standard Borel space and the map c, which inverts the restriction of π to E, is a Borel map of $\mathscr{P}$ into $\mathscr{U}$ such that $\pi(c(x)) = x$ for all $x \in \mathscr{P}$. We now observe that $E_1 = Q_1$ and $1 \in E_1$. Hence we use lemma 7.3 to conclude that $\pi[E_1]$ is open in $\mathscr{P}$ and that c is continuous on $\pi[E_1]$.

2. MULTIPLIERS AND PROJECTIVE REPRESENTATIONS

After these preliminaries involving the projective group, we are in a position to introduce the basic concepts centering around the so-called projective (or ray) representations of a group. We begin with the concept of a multiplier.

Let G, K be lcsc groups with K *abelian*. By a K-*multiplier* for G we mean a Borel map m

$$m : x, y \to m(x,y) \tag{12}$$

of $G \times G$ into K such that

$$\begin{aligned} &\text{(i)} \quad m(x,yz)m(y,z) = m(xy,z)m(x,y) \quad \text{for all} \quad x, y, z \in G, \\ &\text{(ii)} \quad m(x,e) = m(e,x) = 1 \quad \text{for all} \quad x \in G. \end{aligned} \tag{13}$$

Note that we are writing K multiplicatively and that we write 1 for the identity of K. If $K = T$, the multiplicative group of complex numbers of modulus 1, we omit the prefix K and speak simply of *multipliers*. If K is additively written, the equations (13) become

$$\begin{aligned} &\text{(i)} \quad m(xy,z) + m(x,y) = m(x,yz) + m(y,z) \quad \text{for all} \quad x, y, z \in G, \\ &\text{(ii)} \quad m(x,e) = m(e,x) = 0 \quad \text{for all} \quad x \in G. \end{aligned} \tag{14}$$

It follows from (13), on putting $y=x^{-1}$, $z=x$, that

$$m(x,x^{-1}) = m(x^{-1},x) \tag{15}$$

for all $x \in G$.

The set of all K-multipliers for G is obviously a commutative group under pointwise multiplication. We denote this group by $M_K'(G)$. Two K-multipliers m_1, m_2 for G are said to be *similar*, $m_1 \simeq m_2$ in symbols, if there exists a Borel function a on G with values in K,

$$a : x \to a(x)$$

such that

$$m_2(x,y) = \frac{a(xy)}{a(x)a(y)}\, m_1(x,y) \tag{16}$$

for all $x, y \in G$. Note that $a(e)$ is necessarily equal to 1. If a K-multiplier m is similar to the multiplier $\mathbf{1} : x, y \to 1$, we shall say that m is *exact*. Thus m is exact if and only if for some Borel map a of $\mathbf{G}$ into K,

$$m(x,y) = \frac{a(xy)}{a(x)a(y)} \tag{17}$$

for all $x, y \in G$. From (17) it is easily calculated that the exact K-multipliers for G form a subgroup $E_K(G)$ of $M_K'(G)$. We form the quotient group

$$M_K(G) = M_K'(G)/E_K(G) \tag{18}$$

and call it the *K-multiplier group* of G. When $K=T$ we write $M(G)$ for $M_T(G)$ and call it the *multiplier group of G*.

Suppose that m is a multiplier for G. A mapping

$$U : g \to U_g$$

of G into the unitary group $\mathscr{U}$ of a separable Hilbert space $\mathscr{H}$ is said to be an *m-representation* if

$$\begin{array}{ll} \text{(i)} & g \to U_g \quad \text{is Borel, i.e., } g \to \langle U_g f, f' \rangle \quad \text{is Borel for all} \\ & f, f' \in \mathscr{H}, \\ \text{(ii)} & U_e = 1, \\ \text{(iii)} & U_{xy} = m(x,y) U_x U_y \quad \text{for all} \quad x, y \in G. \end{array} \tag{19}$$

A Borel mapping $g \to U_g$ of G into $\mathscr{U}$ is said to be a *projective representation* if there exists a multiplier m for G such that U is an m-representation; then m is obviously uniquely determined by U. We shall say that m is the *multiplier of U*. If U is a projective representation of G in $\mathscr{H}$, the map

$$\pi \circ U : g \to \pi(U_g), \tag{20}$$

of G into the projective group $\mathscr{P}$, is Borel and (iii) of (19) shows that $\pi \circ U$ is a *homomorphism*. Conversely, we shall prove that every Borel homomorphism of G into $\mathscr{P}$ arises from a projective representation in this fashion.

Theorem 7.5. *Let m be any multiplier for G and let U be any m-representation of G acting in $\mathscr{H}$. Then $\pi \circ U$ is a continuous homomorphism of G into $\mathscr{P}$. If V is any projective representation of G in $\mathscr{H}$ such that $\pi \circ U = \pi \circ V$, then the multiplier of V is similar to m; and conversely, if m' is a multiplier similar to m, there is an m'-representation V with $\pi \circ V = \pi \circ U$. Suppose*

$$u : g \to u_g$$

is a Borel homomorphism of G into $\mathscr{P}$. Then there exists a projective representation U of G in $\mathscr{H}$ such that

$$u = \pi \circ U; \tag{21}$$

moreover U can be chosen to be continuous in some open subset of G containing 1. Finally, if m is any multiplier for G, there exist m-representations of G.

Proof. Let U be a projective representation of G in $\mathscr{H}$. Then $x \to \pi(U_x)$ is a Borel map of G into $\mathscr{P}$. Since U_{xy} is a constant multiple of U_xU_y, $\pi(U_{xy}) = \pi(U_x)\pi(U_y)$. Thus $\pi \circ U$ is a homomorphism of G into $\mathscr{P}$. By lemma 5.28, $\pi \circ U$ is continuous. Suppose U' is another projective representation of G in $\mathscr{H}$ and let m and m' be the multipliers of U and U', respectively. Then $\pi \circ U = \pi \circ U'$ if and only if for each $x \in G$ there is an $a(x) \in T$ such that $U_x' = a(x)U_x$. The equation $U_x^{-1}U_x' = a(x)1$ shows that the function $a(x \to a(x))$ is Borel. If such an a exists, then a direct calculation shows that

$$m'(x,y) = m(x,y)\,\frac{a(xy)}{a(x)a(y)} \tag{22}$$

for all $x, y \in G$. Thus $m \simeq m'$. Conversely, suppose $m \simeq m'$. Then (22) is satisfied for a suitable Borel function a on G with values in T and, by setting $U_x' = a(x)U_x$, we obtain an m'-representation U' with $\pi \circ U' = \pi \circ U$.

Next, let

$$u : g \to u_g$$

be a Borel homomorphism of G into $\mathscr{P}$; u is continuous by lemma 5.28. We select a Borel map c of $\mathscr{P}$ into $\mathscr{U}$ satisfying the conditions (i) to (iii) of (7). Let

$$U_g = c(u_g). \tag{23}$$

Since c is Borel and $U=c\circ u$, $U(g\to U_g)$ is a Borel map of G into $\mathscr{U}$. Let $A\subseteq\mathscr{P}$ be an open set containing $\pi(1)$ on which c is continuous. Then U is continuous on the open subset $u^{-1}(A)$ of G containing 1. If $x, y\in G$, then it is obvious that there exists an $m(x,y)\in T$ such that

$$U_{xy} = m(x,y)U_xU_y. \tag{24}$$

Since

$$U_{xy}U_y{}^{-1}U_x{}^{-1} = m(x,y)1,$$

it follows that $x, y\to m(x,y)$ is a Borel map from $G\times G$ into T. The equation $u_xu_{yz}=u_{xy}u_z$, expressed in terms of the U_g's, gives (i) of (13), while the relations $u_{xe}=u_{ex}=u_x$ lead to (ii) of (13). m is, in other words, a multiplier for G, and U is an m-representation, continuous around e.

Finally, let m be an arbitrary multiplier for G. We write $\mathscr{H}=\mathscr{L}^2(G,\mu_r)$, where μ_r is a right Haar measure on G. For any $x\in G$ and $f\in\mathscr{H}$, we define

$$(U_xf)(g) = m(g,x)^{-1}f(gx). \tag{25}$$

It is readily verified that $U_e=1$, $U_{xy}=m(x,y)U_xU_y$. The equation $\|U_xf\|^2=\|f\|^2$ shows that each U_x is unitary. If $f, f'\in\mathscr{H}$,

$$\langle U_xf,f'\rangle = \int_G m(g,x)^{-1}f(gx)f'(g)^*d\mu_r(g),$$

and as the integrand is a Borel function on $G\times X$, $x\to\langle U_xf,f'\rangle$ is a Borel function on G. U is thus an m-representation. The proof of the theorem is complete.

Corollary 7.6. *Let m be any multiplier for G. Then m is similar to a multiplier which is continuous on some open subset of $G\times G$ containing (e,e).*

Proof. Let U be an m-representation of G, acting in a Hilbert space $\mathscr{H}$. From theorem 7.5 we know that there exists a multiplier m' similar to m and an m'-representation U' acting in $\mathscr{H}$ such that (i) $\pi\circ U'=\pi\circ U$, (ii) for a suitable open subset N of G containing e, $x\to U_x'$ is continuous from N into $\mathscr{U}$. Let N_1 be an open set containing e such that $N_1N_1\subseteq N$. As

$$U_{xy}' = m'(x,y)U_x'U_y',$$

and as U' is continuous on N, it is clear that m' is continuous on $N_1\times N_1$.

Corollary 7.7. *Let m be a multiplier for G and U any m-representation of G in $\mathscr{H}$. Then m is exact if and only if there exists a representation U' of G in $\mathscr{H}$ such that $\pi\circ U'=\pi\circ U$. If U'' is another representation of G in $\mathscr{H}$ such that $\pi\circ U''=\pi\circ U$, there exists a continuous homomorphism χ of G into T such that*

$$U_g'' = \chi_gU_g' \tag{26}$$

for all $g\in G$.

Proof. The first assertion is an immediate consequence of theorem 7.5. For the second, we note that $U_g'^{-1}U_g''=\chi_g\cdot 1$ for each $g\in G$, where $\chi_g\in T$. Then $g\to\chi_g$ is Borel and multiplicative. Therefore χ is a continuous homomorphism of G into T.

3. MULTIPLIERS AND GROUP EXTENSIONS

The obvious problem now facing us is of course that of finding, for given lcsc groups G and K (K abelian), the K-multiplier group $M_K(G)$. If, for example, $K=T$ and $M(G)$ is trivial (as it may well be), then we know that every multiplier for G is exact. In this case there is no practical difference between projective and ordinary representations. Our aim is to examine this problem at least for special classes of G and K. We introduce in this section a very useful technique to study the multipliers for a given group, namely the method of group extensions (cf. Mackey [6]).

Suppose G and K are two lcsc groups. We shall, as always, assume that K is abelian. By an *extension of K by G* we mean a triple (H,i,j), where H is a lcsc group, i is an isomorphism of K onto a closed normal subgroup of H, and j is a homomorphism of H onto G with kernel $i[K]$. We require that i be a homeomorphism and that j be continuous. (H,i,j) is said to be a *central* extension if $i[K]$ is contained in the center of H. In the language of cohomology theory, (H,i,j) is an extension of K by G if and only if the sequence

$$0\longrightarrow K\xrightarrow{i}H\xrightarrow{j}G\longrightarrow 0 \tag{27}$$

is exact. If (H,i,j) and (H',i',j') are two extensions, then we say that they are *equivalent*, if there is an isomorphism q of H onto H', such that, the diagram

$$\begin{tikzcd} & H \arrow[dr, "j"] \arrow[dd, "q"] & \\ K \arrow[ur, "i"] \arrow[dr, "i'"'] & & G \\ & H' \arrow[ur, "j'"'] & \end{tikzcd} \tag{28}$$

commutes, i.e., for all $k\in K$,

$$q(i(k)) = i'(k), \tag{29}$$

and for all $h\in H$,

$$j'(q(h)) = j(h). \tag{30}$$

Once again we require that q be a homeomorphism, i.e., q is an isomorphism of the lcsc group H onto the lcsc group H'. We recall that for two lcsc groups A and B, any continuous homomorphism of A into B, which is one-one and maps A onto B, is necessarily a homeomorphism (cf. Pontrjagin [1]).

We shall now associate with any *central* extension (H,i,j) of K by G, a K-multiplier for G. Since H is lcsc and $i[K]$ is a closed subgroup, there is a Borel section c based on G, i.e., a Borel map c of G into H,

$$(31) \qquad c : G \to H$$

such that $c(e) = i(1)$, the identity of H, and for all $x \in G$,

$$j(c(x)) = x$$

(cf. theorem 5.11). Therefore, the map

$$(32) \qquad x, k \to c(x)i(k)$$

is a one-one Borel map of the standard Borel *space* $G \times K$ onto the standard Borel space H. Consequently, it is a Borel isomorphism. We denote by d, the inverse of this map, which is a Borel isomorphism of H onto $G \times K$:

$$(33) \qquad d : c(x)i(k) \to x, k.$$

If $x, y \in G$, $c(xy)$ and $c(x)c(y)$ belong to the same $i[K]$-coset of H, so that

$$(34) \qquad m(x,y) = i^{-1}[c(xy)^{-1}c(x)c(y)]$$

is an element of K. Clearly m is a Borel map of $G \times G$ into K and $m(x,e) = m(e,x) = 1$ for all $x \in G$. If we now remember that $i[m(x,y)]$ commutes with all elements of H, we find, for $x, y, z \in G$,

$$\begin{aligned} i[m(x,yz)m(y,z)] &= c(xyz)^{-1}c(x)c(y)c(z) \\ &= c(xyz)^{-1}c(xy)\{i[m(x,y)]\}c(z) \\ &= c(xyz)^{-1}c(xy)c(z)i[m(x,y)] \\ &= i[m(xy.z)m(x,y)]. \end{aligned}$$

Thus m is a K-multiplier for G. We note at the same time that

$$(35) \qquad c(x)i(k)c(y)i(k') = c(xy)i[m(x,y)kk']$$

for all $x, y \in G$ and $k, k' \in K$. If we use the fact that d is a Borel isomorphism between the Borel spaces H and $G \times K$, then (35) tells us that the definition of a product operation in $G \times K$ given by

$$(36) \qquad (x,k)(y,k') = (xy,m(x,y)kk')$$

converts $G \times K$ into a *Borel group*, and that d provides a Borel group isomorphism of this group with H.

The converse to the result described just now is a much more difficult question. Is it possible to associate with each K-multiplier for G a central extension of K by G? The answer is affirmative as the following theorem, due to Mackey [6], shows.

Theorem 7.8. *Let* m *be any* K*-multiplier. If we define for* (x,k), $(y,k') \in G \times K$,

$$(x,k)(y,k') = (xy, m(xy)kk'), \tag{37}$$

then $G \times K$ *becomes a standard Borel group under this product, and there exists a unique lcsc topology for* $G \times K$ *such that* (a) $G \times K$ *becomes a lcsc group, denoted by* $G \times_m K$, *under this topology, and* (b) *the Borel structure of this topology coincides with the product Borel structure on* $G \times K$. *If*

$$\begin{aligned} i_0(k) &= (e,k), \\ j_0(g,k) &= g, \end{aligned} \tag{38}$$

then $(G \times_m K, i_0, j_0)$ *is a central extension of* K *by* G. *Every central extension of* K *by* G *is equivalent to one such. The extensions corresponding to the* K*-multipliers* m_1 *and* m_2 *are equivalent if and only if* m_1 *and* m_2 *are similar.*

Proof. We begin with a central extension (H,i,j) of K by G and associate with it a K-multiplier m, using the development contained in the equations (32) through (36). Then

$$d : c(x)i(k) \to (x,k)$$

is an isomorphism of H onto $G \times_m K$, where $G \times_m K$ is considered a group with (37) defining the product operation. $G \times_m K$ is obviously a standard Borel group and is also lcsc if we equip it with the topology which makes d a homeomorphism. It is obvious that the extensions (H,i,j) and $(G \times_m K, i_0, j_0)$ are equivalent.

Conversely, let m be a K-multiplier for G. Then under (37), $G \times K$ becomes a standard Borel group, as is readily verified with the help of the relations (13). We write $G \times_m K$ for this group. The relations $m(x,e) = m(e,x) = 1$ show that $\{(e,k) : k \in K\}$ is a central subgroup of $G \times_m K$. Clearly $i_0(k \to (e,k))$ and $j_0((g,k) \to g)$ are Borel homomorphisms. We shall next prove that if λ is a Haar measure on K and μ is a left Haar measure on G, $\mu \times \lambda$ is a left invariant measure on $G \times_m K$. In fact, if f is a $\mu \times \lambda$-summable Borel function on $G \times K$ and $(g_0,k_0) \in G \times K$, we have, by Fubini's theorem,

$$\begin{aligned} \iint_{G \times K} f((g_0,k_0)(g,k))d(\mu \times \lambda)(g,k) &= \iint_{G \times K} f(g_0g, m(g_0,g)k_0k)d(\mu \times \lambda)(g,k) \\ &= \int_G \left(\int_K f(g_0g,k)d\lambda(k)\right)d\mu(g) \\ &= \int_K \left(\int_G f(g_0g,k)d\mu(g)\right)d\lambda(k) \\ &= \int_K \left(\int_G f(g,k)d\mu(g)\right)d\lambda(k) \\ &= \iint_{G \times K} f(g,k)d(\mu \times \lambda)(g,k). \end{aligned}$$

$\mu \times \lambda$ is thus left invariant. Theorem 5.41 now enables us to conclude that $G \times_m K$ becomes a lcsc group under its Weil topology. The Borel structure of this topology is by choice the product Borel structure on $G \times K$, and it is the only lcsc topology with this property, as proved in that theorem. The map

$$j_0 : g, k \rightarrow g$$

is a Borel homomorphism of the lcsc group $G \times_m K$ onto the lcsc group G and is thus continuous. This proves that $i_0[K] = \text{kernel } (j_0)$ is closed in $G \times_m K$. A similar reasoning shows that i_0 is continuous also. Therefore $(G \times_m K, i_0, j_0)$ is a central extension of K by G.

The only thing that remains to be proved is that the extensions $G \times_{m_1} K$ and $G \times_{m_2} K$ are equivalent if and only if $m_1 \simeq m_2$. For any K-multiplier m, let us write

$$H_m = G \times_m K. \tag{39}$$

Suppose $m_1 \simeq m_2$. Then there exists a Borel map a,

$$a : g \rightarrow a(g)$$

of G into K such that

$$m_2(x,y) = m_1(x,y) \frac{a(xy)}{a(x)a(y)}$$

for all $x, y \in G \times G$. If we write

$$q(g,k) = (g,a(g)k), \tag{40}$$

then it is trivial to check that q is a Borel group isomorphism of H_{m_1} onto H_{m_2}. Since H_{m_1} and H_{m_2} are lcsc groups, q is necessarily a homeomorphism. It is obvious that q sets up an equivalence between (H_{m_1}, i_0, j_0) and (H_{m_2}, i_0, j_0). Conversely, let us suppose that H_{m_1} and H_{m_2} are equivalent extensions. Then there exists a Borel isomorphism q of H_{m_1} onto H_{m_2} such that

$$q(e,k) = (e,k)$$

for all $k \in K$, and

$$j_0(q(g,k)) = g$$

for all $g \in G$. Let $a(g) \in K$ be defined by

$$q(g,1) = (g,a(g)). \tag{41}$$

a is clearly a Borel map of G into K. Then, as $(g,1)(e,k) = (g,k)$ we find

$$\begin{aligned} q(g,k) &= (g,a(g))(e,k) \qquad (\text{in } H_{m_2}) \\ &= (g,a(g)k). \end{aligned}$$

If we now use the fact that q preserves the group structures, we can conclude, after a brief computation, that

$$m_2(x,y) = m_1(x,y)\,\frac{a(xy)}{a(x)a(y)}$$

for all $x, y \in G \times G$. In other words, $m_1 \simeq m_2$. The theorem is completely proved.

Remark. The essential difficulty of the theorem proved above resides in the fact that when the K-multiplier m is not continuous on $G \times G$, the product topology would no longer serve to make $G \times_m K$ lcsc. In fact, if m is continuous on $G \times G$, it is obvious that under the product topology, H_m is a topological group and is even a lcsc group. Since the product topology generates the product Borel structure, it follows from the uniqueness of the Weil topology that the two topologies coincide. Conversely, if the Weil topology for H_m coincides with the product topology, the map $x \to (x,1)$ of G into H_m is continuous and hence $x, y \to (x,1)(y,1) = (xy, m(x,y))$ is also continuous, showing that m is continuous. Note that, in this case, $c_0 : x \to (x,1)$ gives a global *continuous* section for $H_m/i_0[K]$, i.e., c_0 is continuous and $j_0(c_0(x)) = x$ for all $x \in G$. This shows that, by using continuous K-multipliers, we can construct only those central extensions (H,i,j) of K by G for which there exists a continuous section c from G into H. There are many examples where such global sections do not exist. For instance, if K is finite but both H and G are compact and *connected*, no continuous section for $H/i[K]$ can be defined on all of G. An example of this is obtained when we take H to be $SU(n,\mathbf{C})$, the special unitary group in n-dimensions, and K as the center of H.

On the other hand, one can impose certain conditions on K under which any K-multiplier for G will be similar to one which is continuous in some open neighborhood of (e,e) in $G \times G$. Corollary 7.6 tells us that this is so when $K = T$. More generally, it can be shown that this is the case when K is a compact Lie group, or when K and G are both Lie groups. We shall call a K-multiplier *locally continuous* if it continues on some open set containing (e,e) in $G \times G$. The assumption of local continuity for m is of a much less serious nature than the assumption of global continuity.

When m is locally continuous, it is almost obvious that one can construct, using local mappings, a fiber space topology for H_m over G which is locally trivial and which coincides with the Weil topology. We shall derive this as a corollary of theorem 7.8. The result is stated in corollary 7.10. We need an auxiliary result.

Corollary 7.9. *Suppose m is locally continuous. Then there exists an open set N around e such that $x \to (x,1)$ is a homeomorphism of N into H_m.*

Proof. We select a symmetric open set A_1 containing e with compact closure such that m is continuous on $\bar{A}_1 \times \bar{A}_1$ and select $A_2, A_3, \cdots$ such that each A_n is open, symmetric, contains e, and $\bar{A}_{n+1}\bar{A}_{n+1} \subseteq A_n$ for all n. Let $\{x_n\}$ be a sequence in $\bar{A}_3$, let $x_0 \in \bar{A}_3$, and let $x_n \to x_0$. We shall show that $(x_n,1) \to (x_0,1)$ in the Weil topology of H_m. Write $y_n = x_0{}^{-1}x_n$. Then $y_n \in A_2$, $y_n \to e$, and

$$(x_0,1)(y_n,1)(e,m(x_0,y_n)^{-1}) = (x_n,1). \tag{42}$$

Since the map $k \to (e,k)$ is continuous from K into H_m and since m is continuous on $\bar{A}_1 \times \bar{A}_1$, (42) shows that $(x_n,1)$ will converge to $(x_0,1)$ provided $(y_n,1) \to (e,1)$.

Consider now $A_2 \times K = j_0{}^{-1}(A_2)$. This is an open subset of H_m containing $(e,1)$ and $(y_n,1)$ belongs to it for all n. In order to prove that $(y_n,1) \to (e,1)$ it is enough to prove that $\|L_{(y_n,1)}f - f\| \to 0$ for all f in $\mathscr{L}^2(H_m)$ which vanish outside $A_2 \times K$. Let D be the set of all functions on $G \times K$, continuous with respect to the *product* topology and vanishing outside subsets of $A_2 \times K$ which are compact in the product topology; it is enough to prove that $L_{(y_n,1)}f \to f$ for all $f \in D$.

Fix $f \in D$ and let $C \subseteq A_2$ and $K_1 \subseteq K$ be compact sets such that f vanishes outside $C \times K_1$. Now,

$$(L_{(y_n,1)}f)(x,k) = f\left(y_n{}^{-1}x, \frac{km(y_n{}^{-1},x)}{m(y_n,y_n{}^{-1})}\right).$$

If $x \notin C$, $y_n{}^{-1}x \notin C$ for sufficiently large n and $(L_{y_n,1}f)(x,k) = f(x,k) = 0$ for sufficiently large n. If $x \in A_2$, the continuity of m on $\bar{A}_1 \times \bar{A}_1$ implies that $(L_{(y_n,1)}f)(x,k) \to f(x,k)$ for all $k \in K$. Now, $|f(x,k)| \leq B$ for all $(x,k) \in G \times K$ and so, $|(L_{(y_n,1)}f)(x,k)| \leq B$ for all n and all (x,k). Further, if $L_1 = m(\bar{A}_1 \times \bar{A}_1)$ then L_1 is compact in K and, for all n, $L_{(y_n,1)}f$ vanishes outside $\bar{A}_1 \times K_1L_1L_1{}^{-1}$ which is compact in the product topology. Hence, by Lebesgue's convergence theorem,

$$\iint\limits_{G \times K} |L_{(y_n,1)}f - f|^2 d(\mu \times \lambda) \to 0. \tag{43}$$

We have thus proved that $x \to (x,1)$ is continuous from $\bar{A}_3$ into H_m. As $\bar{A}_3$ is compact, it is a homeomorphism. This finishes the proof of the corollary.

Corollary 7.10. *If m is continuous around* (e,e), *there exists an open set N containing e such that the Weil topology on the open set $N \times K$ coincides with the product topology.*

Proof. Let N be an open set around e such that $x \to (x,1)$ is a homeomorphism of N into H_m. Since $(x,k) = (x,1)(e,k)$ and since $x \to (x,1)$ and $k \to (e,k)$ are both homeomorphisms, the corollary follows at once.

Remark. Corollary 7.10 shows that when m is locally continuous, H_m, as a fiber space over G, has a locally trivial fibration. For a more intensive study of the topology of H_m we refer the reader to the work of Calabi [1].

The main source of technical complications in our theory is the fact that the topology on H_m in general bears no visible relation to that of G and K. We now describe another corollary to theorem 7.8 which gives an example of the nature of the indirect arguments needed in analyzing H_m.

Corollary 7.11. *If K and G are both connected (resp. simply connected), so is H_m. If K and G are compact, so is H_m.*

Both the assertions follow from the fact that $i_0[K]$ is a closed subgroup of H_m and G is isomorphic to $H_m/i_0[K]$.

Our first use of these group extensions is a description of the similarity classes of K-multipliers. Let m be a K-multiplier for G, and H_m the corresponding group extension. For any Borel map a of G into K with $a(e) = 1$, let

$$\begin{aligned} c_a(x) &= (x,a(x)), \\ m_a(x,y) &= m(x,y)\,\frac{a(x)a(y)}{a(xy)}. \end{aligned} \tag{44}$$

It is easily seen that $m_{a_1} = m_{a_2}$ if and only if there is a Borel homomorphism k of G into K such that $a_1(x) = k(x)a_2(x)$ for all $x \in G$. The map $c_a \to m_a$ establishes a correspondence between multipliers similar to m and Borel sections, and reduces the analysis of the similarity class of m to the analysis of the *geometry* of H_m as a fiber space over G.

Theorem 7.12. *c_a is continuous (respectively continuous around e) if and only if m_a is continuous (respectively continuous around (e,e)). m is exact if and only if there exists a Borel section c which is a homomorphism of G into H_m.*

Proof. Regarding continuity we shall prove only local assertions; the global ones follow from similar reasoning. Suppose now c_a is continuous on an open set N around e. Choose an open set N_1 around e such that $N_1N_1 \subseteq N$. Then the map

$$q : x, y \to c_a(xy)c_a(y)^{-1}c_a(x)^{-1} \tag{45}$$

is continuous from $N_1 \times N_1$ into H_m. A simple calculation yields

$$q(x,y) = (e,m_a(x,y)^{-1}). \tag{46}$$

Since $k \to (e,k)$ is a homeomorphism, (46) shows that m_a is continuous on $N_1 \times N_1$. Conversely, let m_a be continuous on $N \times N$ for a suitable N. If we form the group H_{m_a}, we know from corollary 7.9 that $d : x \to (x,1)$

is a homeomorphism of an open set $N_1 \subseteq N$ containing e into H_{m_a}. Now, from the relation between m and m_a, it follows quickly that

$$p : x, k \to x, ka(x)^{-1}$$

is a Borel group isomorphism of H_m on H_{m_a}. Hence p is a homeomorphism. Therefore, the map $p^{-1} \circ d$ is a homeomorphism of N_1 into H_m. But $(p^{-1} \circ d)(x) = (x,a(x)) = c_a(x)$. Therefore c_a is a homeomorphism of N_1 into H_m.

Now we come to the question of exactness. By very definition, m is exact if and only if for some a, $m_a(x,y) = 1$ for all $x, y \in G$, i.e.,

$$m(x,y) = \frac{a(xy)}{a(x)a(y)} \tag{47}$$

for all $x, y \in G$. A trivial computation shows that (47) is true if and only if $x \to (x,a(x))$ is a *homomorphism* of G into H_m. The proof of theorem 7.12 is complete.

A trivial reformulation of the second half of this theorem gives the following corollary:

Corollary 7.13. *m is exact if and only if there exists a subgroup G' of H_m such that G' is a Borel set meeting each $i_0[K]$-coset exactly once, i.e., $G' \cap i_0[K] = \{(e,1)\}$ and $G' \cdot i_0[K] = H_m$.*

We shall now make a brief digression by obtaining some more corollaries which may serve as illustrations of the use and applicability of the preceding theory.

Corollary 7.14. *Let G be connected and abelian. Then, for a K-multiplier m to be exact, it is necessary and sufficient that $m(x,y) = m(y,x)$ in some neighborhood of (e,e).*

Proof. That an exact multiplier is symmetric (when G is abelian) follows at once from (17). Conversely, let N be an open set containing e such that $m(x,y) = m(y,x)$ for all $x, y \in N \times N$. Then, the formula for multiplication in H_m shows that (x,k) and (y,k') commute whenever x and y belong to N. Since G and K are connected, so is H_m and hence the elements of the form (x,k) $(x \in N)$ generate H_m. This shows that H_m is itself abelian. Since K is isomorphic to $T^r \times V^s$, it follows from the structure of locally compact abelian groups that $\{e\} \times K$ is a direct summand of H_m. Corollary 7.13 now implies that m is exact.

Corollary 7.15. *There exists a multiplier m' similar to m such that m' is locally bounded as a map of $G \times G$ into K (cf. Section 5 of Chapter V).*

Proof. Let a be chosen so that c_a is a *regular* Borel section, i.e., $x \to (x,a(x))$ is a map of G into H_m which sends compact sets of G into sets with compact closure (in H_m). Defining q as in (45), it is obvious that q maps any compact subset of $G \times G$ into a subset of H_m which has compact closure. The formula (46) tells us now that m_a is locally bounded.

We shall end this section with an application to the theory of projective representations.

Theorem 7.16. *Let m be any multiplier for G and let $U(g \rightarrow U_g)$ be an m-representation of G in a (separable) Hilbert space $\mathscr{H}$. Then $(x,k) \rightarrow k^{-1}U_x$ is a representation of H_m in $\mathscr{H}$. Conversely, if $(x,k) \rightarrow V_{(x,k)}$ is a representation of H_m such that $V_{(e,k)} = k^{-1}1$ for all $k \in T$, and if we write $U_x = V_{(x,1)}$, then $U(x \rightarrow U_x)$ is an m-representation of G and $V_{(x,k)} = k^{-1}U_x$.*

Proof. The proof needs only minor calculations and is omitted.

Corollary 7.17. *Let m be any multiplier for G. Then there exists an irreducible m-representation of G.*

Proof. Let V be a representation of H_m such that $V_{(e,k)} = k^{-1}1$ for all $k \in T$. Such V exist, by theorem 7.5, because G admits m-representations. Let us decompose V into a direct integral of irreducible representations V,

$$V \simeq \int V^{\xi} dx(\xi) \tag{48}$$

(cf. Mautner [1], Mackey [5]). Then, from the theory of the direct integrals of representations, it is easy to see that for some ξ_0, $V^{\xi_0}_{(e,k)} = k^{-1}1$ for all $k \in T$ (in fact x-almost all ξ will have this property). Write $V^0 = V^{\xi_0}$, and let U^0 be the corresponding m-representation of G. Since $V^0_{(x,k)} = k^{-1}U_x{}^0$, the irreducibility of V^0 implies that of U^0.

Corollary 7.18. *If G is compact and m is a multiplier for G, then there exist finite dimensional m-representations of G. The irreducible m-representations of G are finite dimensional and every m-representation of G is a direct sum of irreducible m-representations.*

Proof. We observe that H_m is compact (corollary 7.11). The present corollary now follows in exactly the same fashion as the previous one. The proof is even more elementary, since only direct sums, rather than direct integrals, of representations, are involved.

4. MULTIPLIERS FOR LIE GROUPS

We shall now make a deeper study of the group extensions H_m under the assumption that K and G are both connected Lie groups. Since K is abelian, it would be a direct product of a real vector group and a torus.

The first basic result of this section tells us that H_m becomes a Lie group in a natural manner. The way is then opened for the application of results of Lie group theory. We shall use these to show that any K-multiplier for G is similar to one which is globally analytic on $G \times G$ (at least when G is simply connected), and that the elements of the multiplier group are in natural one-one correspondence with equivalence classes of central extensions of the Lie algebra of K by the Lie algebra of G. For general information and terminology involving Lie groups and Lie algebras we refer the reader to the books of Chevalley [1], [2], [3] and Jacobson [2].

Throughout this chapter K will denote a connected abelian Lie group and G a connected Lie group. The symbol T^r will denote a torus of dimension r and the symbol V^s will denote a real vector space of dimension s. Both are regarded as Lie groups in the natural sense.

Our first aim is to introduce an analytic structure on H_m. The essential step in this is accomplished by the following lemma.

Lemma 7.19. *Let K be an additive vector group and m a K-multiplier for G. Then there exists a K multiplier m' similar to m such that m' is C^∞ on $G \times G$.*

Proof. Note that m satisfies the identities (14). It is enough to prove this when K is the additive group of real numbers, since the general case will follow by applying the special case to each component. By virtue of corollary 7.15 we may assume that m is bounded on each compact subset of $G \times G$. Let μ be a left Haar measure and μ_r a right Haar measure for G.

Let then $K = R^1$. Choose a C^∞ function with compact support, say f, which is such that

$$\int_G f(g)d\mu(g) = 1. \tag{49}$$

Let a be the function $x \rightarrow a(x)$, where

$$a(x) = \int_G m(x,u)f(u)d\mu(u). \tag{50}$$

Clearly, a is bounded on each compact set. Let

$$m_1(x,y) = m(x,y) + a(xy) - a(x) - a(y). \tag{51}$$

Then m_1 is a R^1-multiplier similar to m and is bounded on compact subsets of $G \times G$. Moreover,

$$\begin{aligned} m_1(x,y) &= \int_G \{m(x,y) + m(xy,u) - m(y,u) - m(x,u)\}f(u)d\mu(u) \\ &= \int_G \{m(x,yu) - m(x,u)\}f(u)d\mu(u), \end{aligned}$$

using the multiplier identities for m. Thus

$$m_1(x,y) = \int_G m(x,u)f(y^{-1}u)d\mu(u) - \int_G m(x,u)f(u)d\mu(u). \tag{52}$$

The uniform convergence of the integrals on the right of (52) shows that for each $x \in G$, $y \to m_1(x,y)$ is a C^∞ function. Define a_1 to be the Borel function $y \to a_1(y)$, where

$$a_1(y) = \int_G m_1(u,y)f_1(u)d\mu_r(u) \tag{53}$$

f_1 being a C^∞ function with compact support and satisfying

$$\int_G f_1(u)d\mu_r(u) = 1. \tag{54}$$

Let

$$m_2(x,y) = m_1(x,y) + a_1(xy) - a_1(x) - a_1(y). \tag{55}$$

Equation (55) shows that m_2 is a R^1-multiplier similar to m_1 and bounded on compact subsets of $G \times G$. Moreover, a brief calculation shows that

$$m_2(x,y) = \int_G m_1(u,y)f_1(ux^{-1})d\mu_r(u) - \int_G m_1(u,y)f_1(u)d\mu_r(u). \tag{56}$$

From (56) it follows that m_2 is a C^∞ function of each of its variables when the other is fixed. Finally, define a_3 by

$$a_3(x) = \int_G m_2(x,u)f(u)d\mu(u). \tag{57}$$

Write

$$m'(x,y) = m_2(x,y) + a_3(xy) - a_3(x) - a_3(y). \tag{58}$$

Then m' is a R^1-multiplier similar to m and

$$m'(x,y) = \int_G m_2(x,u)f(y^{-1}u)d\mu(u) - \int_G m_2(y,u)f(u)d\mu(u). \tag{59}$$

Since $x, y \to m_2(x,u)f(y^{-1}u)$ is C^∞ on $G \times G$ for each u, (59) shows that m' is C^∞ on $G \times G$. The lemma is proved.

We shall now examine the case when K is not simply connected. We cannot hope for global results but the local version of lemma 7.19 can be proved by the same technique.

Lemma 7.20. *Let K be a connected abelian Lie group and m_0 any K-multiplier. Then there exists a K-multiplier m_0' which is similar to m_0 and which is C^∞ in an open neighborhood of (e,e).*

Proof. Since lemma 7.19 takes care of the case $K=R^s$, it is enough to prove the present result when K is a torus and thus when $K=T$. We may (and do) therefore assume that $K=T$. By corollary 7.6 we may assume that m_0 is locally continuous. It is obvious that we can choose a real bounded Borel function m on $G\times G$ such that

$$(60) \qquad m_0(x,y) = \exp\{im(x,y)\}$$

for all $x, y \in G$, and such that m satisfies the identities characteristic of an R^1-multiplier (cf. (14)) *locally*, i.e., for some symmetric open set M with compact closure containing e,

$$(61) \qquad m(x,e) = m(e,x) = 0$$

for all x in M, and

$$(62) \qquad m(xy,z)+m(x,y) = m(x,yz)+m(y,z)$$

for all $x, y, z \in M$.

We now proceed exactly as we did in the previous lemma. We choose f to be C^∞, and satisfying (49), but with the extra requirement now that f be zero outside a compact subset of M. a is defined as in (50) and m_1 as in (51). Since $\exp(im)=m_0$ is a T-multiplier, (51) shows that $\exp(im_1)$ is a T-multiplier similar to m_0. m_1 is bounded on $G\times G$ and satisfies (61) for $x\in M$ and (62) for $x, y, z\in M$. Equation (52) is now valid for all $x, y \in M$, from which we infer that for each $x\in M$, $y\rightarrow m_1(x,y)$ is C^∞ in M. We define a_1 and m_2 as in (53) and (55), respectively, but after choosing f_1 such that f_1 also vanishes outside M. Once again, (56) is valid for all $x, y\in M$. Proceeding as we did before, we infer finally that m_3 is C^∞ on $M\times M$. This shows that $\exp(im_3)=m'$ is C^∞ on $M\times M$, and is a T-multiplier similar to m_0. The lemma is proved.

We are now in a position to introduce a differentiable structure on H_m. This will then enable us to treat H_m as a Lie group. From this, the introduction of analytic coordinates on H_m will follow in the usual fashion.

Theorem 7.21. *Let K be a connected abelian Lie group and G a connected Lie group. Let m be any K-multiplier for G. Then there exists a unique analytic structure for H_m (compatible with the Weil topology) which converts H_m into an analytic group. The maps $k\rightarrow(e,k)$, of K into H_m, and $(g,k)\rightarrow g$, of H_m onto G, are analytic.*

Proof. H_m is connected by corollary 7.11. We know that any topological group admits at most one compatible analytic structure which converts it into a Lie group (cf. Chevalley [1], pp. 127–129). Thus we are concerned only with existence. Since we may pass to K-multipliers similar to m, we may assume, after lemma 7.20, that m itself is C^∞ around (e,e). From corollary 7.10 we know that for some open set M containing e, the Weil topology on $M \times K$ coincides with the product topology. We choose M so small that this condition is satisfied and that m is C^∞ on $M \times M$.

We give to $M \times K$ the product of the C^∞ structures on M and K and choose a symmetric open set M_1 containing e such that $M_1 M_1 \subseteq M$. For (x,k) and $(y,k') \in M_1 \times K$,

$$(x,k)(y,k')^{-1} = \left(xy^{-1}, kk'^{-1} \frac{m(x,y^{-1})}{m(y,y^{-1})}\right),$$

which shows that the map $(x,k), (y,k') \rightarrow (x,k)(y,k')^{-1}$ is C^∞ on $(M_1 \times K) \times (M_1 \times K)$. By a standard result in Lie group theory (cf. Chevalley [1], pp. 134–135) there exists a C^∞ structure on H_m converting H_m into a Lie group, and such that the induced C^∞ structure on $M \times K$ coincides with the product C^∞ structure around $(e,1)$. In this manner one converts H_m into a connected Lie group. Note, under the assumption that m is C^∞ around (e,e), that there exist open sets $M_2 \subseteq G$ and $K_2 \subseteq K$ containing the respective identities, such that $M_2 \times K_2$ has the product C^∞ structure.

Now, it is known (cf. Pontrjagin [1]) that every C^∞ Lie group can be given a unique compatible analytic structure. H_m may be given this unique analytic structure. The final statement of the theorem follows from the known fact (cf. Chevalley [1], pp. 128–129) that a continuous homomorphism of one analytic group into another is necessarily analytic. The proof of the theorem is complete.

We can deduce several corollaries from this theorem.

Corollary 7.22. *The map $x \rightarrow (x,1)$ of G into H_m is C^∞ around e, m being locally C^∞.*

This follows from the fact that the C^∞ structure of H_m is, around $(e,1)$, the product structure.

Corollary 7.23. *Let m be an arbitrary K-multiplier for G. Then there exists a multiplier m' which is similar to m and is C^∞ (respectively analytic) on all of $G \times G$ if and only if there exists a C^∞ (respectively analytic) map $x \rightarrow c(x)$ of G into H_m such that $j_0(c(x)) = x$ for all x, i.e., if and only if there exists a global C^∞ (respectively analytic) section.*

Proof. The proof is the same as for theorem 7.12; we need only replace "continuous" by "C^∞" and "analytic" at the appropriate places.

Corollary 7.24. *Given any K-multiplier for G, there exists a multiplier similar to it which is analytic around (e,e).*

Proof. This requires the result that, for any analytic group, the coset space determined by a closed subgroup admits a local analytic section (cf. Chevalley [1], pp. 109–111). Applying this result to the analytic group H_m and the closed subgroup $\{e\} \times K = i_0[K]$, we find a Borel section c,

$$c : x \to (x,a(x))$$

such that for some open set N around e, c is an analytic map on $N \cdot N$. If we define

$$q(x,y) = c(xy)c(y)^{-1}c(x)^{-1},$$

then $q(x,y) = (e,m_a(x,y)^{-1})$ (cf. (46)) and q is an analytic map of $N \times N$ into H_m. Since $k \to (e,k)$ is an analytic isomorphism and since $\{e\} \times K$ is a closed analytic submanifold of H_m, we deduce that m_a is analytic on $N \times N$. Obviously $m_a \simeq m$.

This is as far as we can go with elementary methods. To obtain further, and in particular, global results, greater use must be made of Lie group theory. Our aim is to use some of the ideas of Hochschild [1], which in turn are based on a very interesting lemma of Malčev [1], to prove that when G is connected and simply connected, there exists a canonical one-one correspondence between the Lie group extensions and the associated Lie algebra extensions. We shall moreover be able to show that analytic multipliers exist in every similarity class and that they are naturally induced by analytic K^*-multipliers for G, K^* being a covering group of K.

We begin with some preliminary remarks. For any analytic homomorphism a of a Lie group A into a Lie group B, we write $\dot{a}$ for the induced homomorphism of the Lie algebras; $a(\exp X) = \exp \dot{a}(X)$ for $X \in \mathfrak{a}$, where $\mathfrak{a}$ is the Lie algebra of A. Let $\mathfrak{b}$ be the Lie algebra of B and let us assume that A and B are both connected and simply connected. Consider now a homomorphism $d(X \to d(X))$ of the Lie algebra $\mathfrak{b}$ into the Lie algebra of all endomorphisms of the vector space $\mathfrak{a}$ such that $d(X)$ is a *derivation* of $\mathfrak{a}$ for all X, i.e., $d(X)([Y,Z]) = [d(X)(Y),Z] + [Y,d(X)(Z)]$ for all Y, Z in $\mathfrak{a}$. Then there exists a unique analytic homomorphism $b \to \delta(b)$ of B into the Lie group of invertible linear transformations of the vector space $\mathfrak{a}$ such that $\delta(\exp X) = \exp d(X)$ for all $X \in \mathfrak{b}$. Since $d(X)$ is a derivation of $\mathfrak{a}$, $\delta(\exp X)$ is an automorphism of the Lie algebra $\mathfrak{a}$. Hence $\delta(b)$ is an automorphism of $\mathfrak{a}$ for all b. The map $b, Y \to \delta(b)Y$ is clearly analytic from $B \times \mathfrak{a}$ to $\mathfrak{a}$ ($\mathfrak{a}$ being a finite dimensional real vector space, can be considered in an obvious fashion as a real analytic manifold). Since A is simply connected, each $\delta(b)$ gives rise to a unique automorphism $D(b)$ of A such that $D(b)(\exp Y) = \exp\{\delta(b)(Y)\}$ for all $Y \in \mathfrak{a}$. Since $b \to \delta(b)$ is a homomorphism, $b \to D(b)$ is a homomorphism of B into the group of all analytic automorphisms of the Lie group A. We thus see that A becomes a B-space. Moreover, the equation $D(b)(\exp Y) = \exp\{\delta(b)Y\}$ shows that the map $b, a \to D(b)a$ is an analytic map of $B \times A$ into A. Therefore if we form the semidirect product $B \times_D A$ (cf. Chapter VI),

$B \times_D A$ becomes a connected, simply connected Lie group. Its analytic structure is just the product structure of $B \times A$. As usual, we shall identify A and B with subgroups of $B \times_D A$; A will be a closed normal subgroup, and for $b \in B$, $a \in A$, bab^{-1} will coincide with $D(b)a$. Moreover, for $Y \in \mathfrak{a}$ and $X \in \mathfrak{b}$, $\exp X \exp Y \exp(-X) = \exp\{\exp(d(X))Y\}$. This shows that, if we identify the Lie algebra of $B \times_D A$ with the direct sum $\mathfrak{b} + \mathfrak{a}$, then for $Y \in \mathfrak{a}$ and $X \in \mathfrak{b}$, $[X, Y] = d(X)Y$.

Lemma 7.25. *Let* $\mathfrak{g}$ *be a Lie algebra over a field of characteristic zero and* $\mathfrak{g}_1$ *an ideal in* $\mathfrak{g}$. *Suppose that either* $\dim(\mathfrak{g}/\mathfrak{g}_1) = 1$ *or* $\mathfrak{g}/\mathfrak{g}_1$ *is semisimple. Then* $\mathfrak{g}_1$ *is a semidirect summand of* $\mathfrak{g}$, *i.e., there exists a Lie subalgebra* $\mathfrak{p}$ *such that* $\mathfrak{g} = \mathfrak{g}_1 + \mathfrak{p}$ *and* $\mathfrak{g}_1 \cap \mathfrak{p} = 0$.

Proof. If $\dim(\mathfrak{g}/\mathfrak{g}_1) = 1$, then we may take for $\mathfrak{p}$ any one-dimensional subspace not in $\mathfrak{g}_1$. In case $\mathfrak{g}/\mathfrak{g}_1$ is semisimple, the assertion is a refinement of the well known Levi-Malčev theorem and follows easily from it. Let $\mathfrak{r}$ be the radical of $\mathfrak{g}$. Let π be the canonical map of $\mathfrak{g}$ on $\mathfrak{g}/\mathfrak{g}_1$. Since $\pi[\mathfrak{r}]$ is a solvable ideal of $\mathfrak{g}/\mathfrak{g}_1$, $\pi[\mathfrak{r}] = 0$, so that $\mathfrak{r} \subseteq \mathfrak{g}_1$. Let $\mathfrak{p}'$ be a semisimple subalgebra of $\mathfrak{g}$ such that $\mathfrak{p}' \cap \mathfrak{r} = 0$ and $\mathfrak{p}' + \mathfrak{r} = \mathfrak{g}$. Such a $\mathfrak{p}'$ exists, by virtue of the Levi-Malčev theorem (cf. Chevalley [3], pp. 135–145). As $\mathfrak{p}' \cap \mathfrak{g}_1$ is an ideal in $\mathfrak{p}'$, there exists a subalgebra $\mathfrak{p}$ (which is even an ideal in $\mathfrak{p}'$) such that $\mathfrak{p} + (\mathfrak{p}' \cap \mathfrak{g}_1) = \mathfrak{p}'$ and $\mathfrak{p} \cap (\mathfrak{p}' \cap \mathfrak{g}_1) = 0$ (cf. Chevalley [3], pp. 64–65). $\mathfrak{p}$ is semisimple and it is trivial to verify that $\mathfrak{p} \cap \mathfrak{g}_1 = 0$ and $\mathfrak{p} + \mathfrak{g}_1 = \mathfrak{g}$.

Lemma 7.26 (*Malčev* [1]). *Let* H *be a connected and simply connected Lie group and* K *a closed connected normal subgroup of* H. *Let* γ *be the canonical homomorphism of* H *onto* H/K. *Then there exists an analytic map* c *of* H/K *into* H *such that*

$$\gamma(c(t)) = t \tag{63}$$

for all $t \in H/K$.

Proof. We follow Hochschild's proof [1]. Let $\mathfrak{h}$ be the Lie algebra of H and $\mathfrak{k}$ the Lie subalgebra which corresponds to K. Since K is normal, $\mathfrak{k}$ is an ideal in $\mathfrak{h}$. Let $\mathfrak{h} = \mathfrak{h}_1 \supseteq \mathfrak{h}_2 \supseteq \cdots \supseteq \mathfrak{h}_r = \mathfrak{k}$ be a composition series of $\mathfrak{h}$ over $\mathfrak{k}$, i.e., each $\mathfrak{h}_{i+1}$ is an ideal in $\mathfrak{h}_i$, and is maximal in $\mathfrak{h}_i$. We prove lemma 7.26 by induction on r, the length of the composition series.

Since $\mathfrak{h}_2$ is maximal in $\mathfrak{h}_1$, either $\dim(\mathfrak{h}_1/\mathfrak{h}_2) = 1$ or $\mathfrak{h}_1/\mathfrak{h}_2$ is semisimple. Therefore, by lemma 7.25, there exists a subalgebra $\mathfrak{b}$ of $\mathfrak{h}_1$ such that $\mathfrak{h}_1 = \mathfrak{b} + \mathfrak{h}_2$ and $\mathfrak{b} \cap \mathfrak{h}_2 = 0$. For $X \in \mathfrak{b}$ and $Y \in \mathfrak{h}_2$, write

$$d(X)Y = [X, Y]. \tag{64}$$

Let H_2 and B be connected simply connected groups corresponding to $\mathfrak{h}_2$ and $\mathfrak{b}$, respectively. In view of our earlier remarks, the map $X \to d(X)$

gives rise to an analytic map $b, a \rightarrow D(b)a$ of $B \times H_2$ into H_2 such that H_2 becomes a B-space and for each $b \in B$, $D(b)$ is an analytic automorphism of H_2. We may therefore form the group $B \times_D H_2$ and identify its Lie algebra with $\mathfrak{h}_1$ in such a way that $B \times \{e_{H_2}\}$ and $\{e_B\} \times H_2$ correspond to $\mathfrak{b}$ and $\mathfrak{h}_2$, respectively (e_B is the identity of B and e_{H_2} that of H_2). Now, $B \times_D H_2$ is simply connected and so is H. Hence we may identify H with $B \times_D H_2$. In other words, we may assume that H_2 and B are the analytic subgroups of H corresponding to $\mathfrak{h}_2$ and $\mathfrak{b}$, respectively. Then $H_2 \cap B = \{e_H\}$ (the identity of H) and $H = H_2 B$.

Let $\overline{H} = H/K$ and let γ be the canonical homomorphism of H onto $\overline{H}$. Put $\overline{H}_2 = \gamma(H_2), \overline{B} = \gamma(B)$. Now, for each $b \in B$, $D(b)$ leaves K and H_2 invariant and so induces an automorphism $\overline{D}(b)$ of H_2/K. Clearly, $b, a \rightarrow \overline{D}(b)a$ is analytic from $B \times (H_2/K)$ into H_2/K and so one can form $B \times_{\overline{D}} (H_2/K)$. It is obvious that our identification of H with $B \times_D H_2$, gives rise to a natural identification of $\overline{H}$ with $B \times_{\overline{D}} (H_2/K)$. This proves that $\overline{H}_2$ and $\overline{B}$ are closed in $\overline{H}$, that $\overline{H}_2 \cap \overline{B} = \{e_{\overline{H}}\}$ and $\overline{H} = \overline{H}_2 \overline{B}$. In particular, γ is an isomorphism of B onto $\overline{B}$. By the induction hypothesis, there is an analytic map c' ($\bar{s} \rightarrow c'(\bar{s})$) of $\overline{H}_2$ into H_2 such that $\gamma(c'(\bar{s})) = \bar{s}$ for all $\bar{s} \in \overline{H}_2$. Define c on $\overline{H}$ by

$$c(\bar{s}\gamma(b)) = c'(\bar{s})b \qquad (\bar{s} \in \overline{H}_2, b \in B). \tag{65}$$

Then c satisfies (63).

Corollary 7.27. *The map*

$$k, t \rightarrow c(t)k \tag{66}$$

is an analytic isomorphism of the manifold $K \times (H/K)$ onto the manifold H.

Proof. Let ζ denote this map. ζ is analytic and one-one. For $a \in H$,

$$\zeta^{-1}(a) = (c(\gamma(a))^{-1}a, \gamma(a))$$

which is clearly analytic.

Corollary 7.28. *Let $\mathfrak{h}$ be a Lie algebra over the reals, H a connected and simply connected Lie group corresponding to $\mathfrak{h}$. Let $\mathfrak{k}$ be an ideal in $\mathfrak{h}$. Then the analytic subgroup K of H which corresponds to $\mathfrak{k}$ is closed. Moreover both K and H/K are simply connected.*

Proof. That K is closed is well known (cf. Chevalley [1], pp. 125–127). Corollary 7.27 implies that H is (topologically) homeomorphic to $K \times H/K$, and so the simple connectivity of H implies that of K and H/K.

Remark. It is known that if K and H/K are simply connected, so is H (cf. Chevalley [1], pp. 59–60). Thus corollary 7.28, together with this result, completely solves the problem of simple connectivity for H, K, and H/K, when K is closed, connected, and normal.

Lemma 7.29 (*Hochschild* [1]). *Let G be a connected, simply connected Lie group and (H,i,j) an extension of the connected Lie group K by G. Then there exists an analytic map c of G into H such that*

$$j(c(g)) = g \tag{67}$$

for all $g \in G$.

Proof. Let H^* be a connected and simply connected covering group of H and let δ be the covering homomorphism of H^* onto H. Then $j \circ \delta$ is a homomorphism of H^* onto G. Let K_1 be the kernel of $j \circ \delta$ and $K_1{}^0$ the component of the identity of K_1. If $K_1{}^0 \neq K_1$, $H^*/K_1{}^0$ would provide a nontrivial covering for G, which is not possible since G is simply connected. Hence $K_1 = K_1{}^0$, i.e., K_1 itself is connected. From Malčev's lemma we know that there exists an analytic map c^* of G into H^* such that

$$(j \circ \delta)(c^*(g)) = g$$

for all $g \in G$. If we write $c = \delta \circ c^*$, then c is analytic from G into H and satisfies (67) for all $g \in G$.

Corollary 7.30. *Let G be a connected, simply connected Lie group and K a connected abelian Lie group. If m is a K-multiplier for G, then there exists a K-multiplier m' similar to m which is analytic on $G \times G$.*

Proof. Let H_m be the extension associated with m. From lemma 10.29 we know that there exists a global analytic section from G into H_m. Then, corollary 7.23 implies that there is a multiplier similar to m which is globally analytic over $G \times G$.

Remark. That m' can be chosen to be C^∞ was proved by Bargmann [1].

We shall use the lemmas of Hochschild and Malčev to obtain a direct relation between central extensions of Lie groups and the corresponding central extensions of Lie algebras. Let $\mathfrak{k}$, $\mathfrak{g}$ be Lie algebras over $\mathbf{R}$ with $\mathfrak{k}$ abelian. A central extension of $\mathfrak{k}$ by $\mathfrak{g}$ is a triple $(\mathfrak{h},\alpha,\beta)$, where $\mathfrak{h}$ is a Lie algebra over $\mathbf{R}$, α is an isomorphism of $\mathfrak{k}$ into the center of $\mathfrak{h}$, and β a homomorphism of $\mathfrak{h}$ onto $\mathfrak{g}$ such that $\alpha[\mathfrak{k}] = \text{kernel}\ (\beta)$, i.e.,

$$[\alpha[\mathfrak{k}],\mathfrak{h}] = 0 \tag{68}$$

and

$$0 \longrightarrow \mathfrak{k} \xrightarrow{\alpha} \mathfrak{h} \xrightarrow{\beta} \mathfrak{g} \longrightarrow 0 \tag{69}$$

is an exact sequence. Central extensions $(\mathfrak{h},\alpha,\beta)$ and $(\mathfrak{h}',\alpha',\beta')$ of $\mathfrak{k}$ by $\mathfrak{g}$ are said to be equivalent if there exists a (Lie) isomorphism ξ of $\mathfrak{h}$ onto $\mathfrak{h}'$ such that the diagram

$$\begin{tikzcd}
\mathfrak{k} \arrow[r, "\alpha"] \arrow[dr, "\alpha'"'] & \mathfrak{h} \arrow[r, "\beta"] \arrow[d, "\xi"] & \mathfrak{g} \\
& \mathfrak{h}' \arrow[ur, "\beta'"'] &
\end{tikzcd} \tag{70}$$

commutes. Suppose K is a connected abelian Lie group, G a connected Lie group, and (H,i,j) a central extension of K by G. Let $\mathfrak{k}$, $\mathfrak{h}$, and $\mathfrak{g}$ be the Lie algebras of K, H, and G, respectively. Then it is obvious that $(\mathfrak{h},\dot{i},\dot{j})$ is a central extension of $\mathfrak{k}$ by $\mathfrak{g}$. We shall say that $(\mathfrak{h},\dot{i},\dot{j})$ is the *associated central extension* of $\mathfrak{k}$ by $\mathfrak{g}$. A routine transport of structure shows that corresponding to equivalent central extensions of K by G, are associated equivalent central extensions of $\mathfrak{k}$ by $\mathfrak{g}$.

Theorem 7.31. *Let G, K be connected Lie groups and $\mathfrak{g}$, $\mathfrak{k}$ the corresponding Lie algebras. Suppose that K is abelian and G is simply connected. For any equivalence class τ of central extensions of K by G, let $\dot{\tau}$ denote the equivalence class of the associated central extensions of $\mathfrak{k}$ by $\mathfrak{g}$. Then $\tau \to \dot{\tau}$ is a one-one correspondence and its range is the set of all equivalence classes of central extensions of $\mathfrak{k}$ by $\mathfrak{g}$.*

Proof. We shall first give consideration to the case when K is also simply connected. Let $(\mathfrak{h},\alpha,\beta)$ be any central extension of $\mathfrak{k}$ by $\mathfrak{g}$. Let H be a connected and simply connected Lie group which corresponds to $\mathfrak{h}$. Since K is simply connected, there exists an analytic homomorphism i of K into H such that $\dot{i}=\alpha$; similarly there exists an analytic homomorphism j of H onto G such that $\dot{j}=\beta$. The range $i[K]$ being an analytic subgroup of H having the ideal $\alpha[\mathfrak{k}]$ for its Lie algebra, corollary 7.28 implies that $i[K]$ is closed and simply connected. Hence i is also an isomorphism. On the other hand, $i[K]$ is the connected component containing the identity of the kernel K_1 of j. If $i[K]\neq K_1$, $H/i[K]$ will be a nontrivial covering of $G\simeq H/K_1$, which is impossible as G is simply connected. Therefore $i[K]$ is precisely the kernel of j. In other words, (H,i,j) is a central extension of K by G and $(\mathfrak{h},\alpha,\beta)$ is its associated extension. The uniqueness of i and j, once H is specified, implies at once that two extensions of K by G, constructed in this manner, are equivalent if and only if their associated Lie algebra extensions are equivalent. Thus we have proved the lemma in the special case when K is simply connected.

Suppose now that K is connected but not simply connected. Let K^* be a connected, simply connected Lie group covering K. Let ε be the covering homomorphism. Let Z be the kernel of ε. In view of the result for the simply connected case, it is enough to set up a natural correspondence between central extensions of K and K^* (by G) which respects equivalence.

Let (H^*,i^*,j^*) be a central extension of K^* by G. Then H^* is simply connected. Let $H = H^*/i^*Z$ and let δ be the canonical homomorphism of H^* onto H. Then δ has discrete kernel. Since kernel $(\delta) \subseteq i^*[K^*]$, there is a unique homomorphism j of H onto G such that $j^* = j \circ \delta$. Since kernel $(\delta) = i^*$ [kernel (ε)] there is a unique homomorphism i of K into H such that $\delta \circ i^* = i \circ \varepsilon$; i is then injective. It is then obvious that kernel $(j) = i[K]$. Thus the diagram

(71)
$$\begin{tikzcd}
H^* \arrow[rrrr, "\delta"] \arrow[ddrr, "j^*"'] & & & & H \arrow[ddll, "j"] \\
& K^* \arrow[ul, "i^*"'] \arrow[rr, "\varepsilon"'] & & K \arrow[ur, "i"] & \\
& & G & &
\end{tikzcd}$$

commutes. Moreover, (H,i,j) is a central extension of K by G.

Suppose conversely that (H,i,j) is a central extension of K by G. Let H^* be a connected, simply connected Lie group which covers H. Let δ be the covering homomorphism. Then there exists a unique homomorphism i^* of K^* into H^* such that $\delta \circ i^* = i \circ \varepsilon$. Write $j^* = j \circ \delta$. As G is simply connected, we conclude that kernel (j^*) is connected. From this it follows quickly that kernel $(j^*) = i^*[K^*]$. Hence $i^*[K^*]$ is simply connected. Since K^* is also simply connected, i^* is injective. Hence (H^*,i^*,j^*) is a central extension of K^* by G. Moreover, the diagram (71) commutes, and

(72)
$$\text{kernel } (\delta) = i^*[Z].$$

We have thus set up a natural correspondence between central extensions of K and K^* by G. Routine diagram chasing now shows that this correspondence respects equivalence. The proof of theorem 7.31 is complete.

Let (H^*,i^*,j^*) be a central extension of K^* by G and let (H,i,j) be a central extension of K by G. Suppose there is a homomorphism δ of H^* onto H such that (71) is a commutative diagram. Then we say that (H^*,i^*,j^*) *covers* (*through* ε) (H,i,j).

Corollary 7.32. *Let G be connected and simply connected. Let K, K^* and ε be as above. Then the map $m^* \to \varepsilon \circ m^*$ maps the set of all analytic (respectively C^∞) K^*-multipliers onto the set of all analytic (respectively C^∞)*

K-multipliers. $m_1^* \simeq m_2^*$ if and only if $\varepsilon \circ m_1^* \simeq \varepsilon \circ m_2^*$ so that $m^* \to \varepsilon \circ m^*$ induces an isomorphism of $M_{K^*}(G)$ onto $M_K(G)$.

Proof. Let m be an analytic K-multiplier for G. Let H_m be the associated extension. Then H_m has the product analytic structure. Since we can cover the extension H_m by an extension of K^* by G, we may assume that for some K^*-multiplier m_1^*, the extension $H_{m_1^*}$ covers H_m. In view of corollary 7.30, we may assume that m_1^* is analytic on $G \times G$ so that $H_{m_1^*}$ has the product analytic structure.

Let $S = \{(g,1) : g \in G\} \subseteq H_m$. Let δ be the covering homomorphism of $H_{m_1^*}$ onto H_m. Write $S^* = \delta^{-1}(S)$. A standard topological argument shows that δ is a covering map of each connected component of S^* onto S. Let S_0 be the component of S^* containing the identity $(e,1)$ of $H_{m_1^*}$. Then S_0 is a closed regular analytic submanifold of $H_{m_1^*}$. Since S is simply connected, δ is an analytic isomorphism of S_0 onto S_1. Hence there is an analytic map a of G into K^* such that for all $g \in G$,

$$\delta((g,a(g))) = (g,1). \tag{73}$$

Since δ is a homomorphism, (73) gives

$$\delta(xy,m_1^*(x,y)a(x)a(y)) = \delta((xy,a(xy)))(e,m(x,y)). \tag{74}$$

Now, $(xy,a(xy))^{-1}(xy,a(x)a(y)m_1^*(x,y))$ is computed to be

$$\left(e, \frac{a(x)a(y)}{a(xy)}\, m_1^*(x,y)\right)$$

so that (74) leads to the equation

$$\delta\left(e, \frac{a(x)a(y)}{a(xy)}\, m_1^*(x,y)\right) = (e,m(x,y)),$$

and consequently,

$$\varepsilon\left(\frac{a(x)a(y)}{a(xy)}\, m_1^*(x,y)\right) = m(x,y) \tag{75}$$

for all $x, y \in G$. If we write $m^*(x,y)$ for the element in (75) to which ε is applied, we have

$$m = \varepsilon \circ m^* \tag{76}$$

m^* is clearly analytic. This proves the result in the analytic case. The C^∞ case can be disposed off similarly.

The equation $\varepsilon \circ m^* = m$ implies, on the other hand, that the map $x, k^* \to x, \varepsilon(k^*)$ is a homomorphism of H_{m^*} on H_m which sets up a covering of the extension H_m by H_{m^*}. From theorem 7.31 we now conclude immediately that $m_1^* \simeq m_2^*$ if and only if $\varepsilon \circ m_1^* \simeq \varepsilon \circ m_2^*$. From this it follows at once that $m^* \to \varepsilon \circ m^*$ induces an isomorphism of $M_{K^*}(G)$ onto $M_K(G)$.

Corollary 7.33. *Let G be a connected, simply connected Lie group and K a connected abelian Lie group. If m is a K-multiplier, then m is exact if and only if, in the extension of $\mathfrak{k}$ by $\mathfrak{g}$ which is associated with H_m, the image of $\mathfrak{k}$ is a semidirect summand.*

Proof. Let $(\mathfrak{h},\alpha,\beta)$ be the extension of $\mathfrak{k}$ by $\mathfrak{g}$ which is associated with (H_m,i_0,j_0) (where $i_0(k)=(e,k)$ and $j_0(g,k)=g$). If m is exact, there exists a Borel homomorphism $c(x \to c(x))$ of G into H_m such that $j_0(c(x))=x$ for all $x \in G$ (theorem 7.12). Since G and H_m are Lie groups, c is analytic. Hence the range B of c is an analytic subgroup of H_m. If $\mathfrak{b}$ is the subalgebra of $\mathfrak{h}$ corresponding to B, it is obvious that $\mathfrak{b} \cap \alpha[\mathfrak{k}]=0$ and $\mathfrak{b}+\alpha[\mathfrak{k}]=\mathfrak{h}$.

Conversely, suppose there exists a subalgebra $\mathfrak{b}$ of $\mathfrak{h}$ such that $\mathfrak{b}+\alpha[\mathfrak{k}]=\mathfrak{h}$ and $\mathfrak{b} \cap \alpha[\mathfrak{k}]=0$. Then there exists obviously an isomorphism γ of $\mathfrak{g}$ into $\mathfrak{h}$ such that γ maps $\mathfrak{g}$ onto $\mathfrak{b}$ and $\beta \circ \gamma$ is the identity. Since G is simply connected, there is a homomorphism c of G into H_m such that $\dot{c}=\gamma$. Since $\beta \circ \gamma$ is the identity, $j_0 \circ c$ is locally the identity. Since G is connected, $j_0 \circ c$ is the identity everywhere. This proves that m is exact.

The only problem which is still to be examined is that of classifying the central extensions of $\mathfrak{k}$ by $\mathfrak{g}$. This can be done very simply as the problem is *linear*. We now proceed to indicate the standard solution to this problem (cf. Chevalley-Eilenberg [1]). Let p be a bilinear map of $\mathfrak{g} \times \mathfrak{g}$ into $\mathfrak{k}$ such that (i) p is skew symmetric, i.e.,

$$p(X,Y)+p(Y,X) = 0 \tag{77}$$

for all X, $Y \in \mathfrak{g}$, and (ii) p satisfies the identity

$$p([X,Y],Z)+p([Y,Z],X)+p([Z,X],Y) = 0 \tag{78}$$

for all X, Y, $Z \in \mathfrak{g}$. We shall associate with p a central extension of $\mathfrak{k}$ by $\mathfrak{g}$. Let

$$\mathfrak{h}_p = \mathfrak{g} \times \mathfrak{k},$$

and let us define, for (X,S) and $(X',S') \in \mathfrak{h}_p$,

$$[(X,S),(X',S')] = ([X,X'],p(X,X')). \tag{79}$$

From the identities (77) and (78) it follows that $\mathfrak{h}_p$ becomes a Lie algebra under the bracket operation (79). Moreover, (79) shows trivially that $(X,S) \to X$ is a homomorphism of $\mathfrak{h}_p$ onto $\mathfrak{g}$, that $(0,S)$ lies in the center of $\mathfrak{h}_p$ for all $S \in \mathfrak{k}$, and that $S \to (0,S)$ is an isomorphism of $\mathfrak{k}$ into $\mathfrak{h}_p$. Write

$$\begin{aligned} \alpha_0(S) &= (0,S), \\ \beta_0(X,S) &= X. \end{aligned} \tag{80}$$

$(\mathfrak{h}_p,\alpha_0,\beta_0)$ is then a central extension of $\mathfrak{k}$ by $\mathfrak{g}$.

Lemma 7.34. *Every central extension of* $\mathfrak{k}$ *by* $\mathfrak{g}$ *is equivalent to an extension* $\mathfrak{h}_p$ *for a suitable bilinear map* p *of* $\mathfrak{g}\times\mathfrak{g}$ *into* $\mathfrak{k}$ *satisfying* (77) *and* (78). $\mathfrak{h}_{p_1}$ *and* $\mathfrak{h}_{p_2}$ *are equivalent extensions if and only if there exists a linear map* q *of* $\mathfrak{g}$ *into* $\mathfrak{k}$ *such that*

$$p_2(X,Y)-p_1(X,Y) = q([X,Y]) \tag{81}$$

for all X, $Y \in \mathfrak{g}$.

Proof. Let $(\mathfrak{h},\alpha,\beta)$ be a central extension of $\mathfrak{k}$ by $\mathfrak{g}$. By identifying $\mathfrak{k}$ with $\alpha[\mathfrak{k}]$ and $\mathfrak{g}$ with a linear manifold supplementary to $\alpha[\mathfrak{k}]$, we may assume that $\mathfrak{h}=\mathfrak{g}\times\mathfrak{k}$, $\alpha=\alpha_0$, and $\beta=\beta_0$. We have now only to determine the form of $[.,.]$ for pairs of elements of $\mathfrak{h}$. Since $(X,S),(X',S') \to [(X,S),(X',S')]$ is bilinear and skew symmetric and since $(X,S) \to X$ is a homomorphism of $\mathfrak{h}$ onto $\mathfrak{g}$, it follows that there exists a skew symmetric bilinear map of $\mathfrak{h}\times\mathfrak{h}$ into $\mathfrak{k}$, say r, such that

$$[(X,S),(X',S')] = ([X,X'],r(X,S;\, X',S')) \tag{82}$$

for all X, $X' \in \mathfrak{g}$ and S, $S' \in \mathfrak{k}$. The fact that $(0,S)$ lies in the center of $\mathfrak{h}$ for all S in $\mathfrak{k}$ now implies that

$$r(0,S;\, X',S') = 0 \tag{83}$$

for all $X' \in \mathfrak{g}$ and S, $S' \in \mathfrak{k}$. The bilinear maps of $\mathfrak{h}\times\mathfrak{h}$ into $\mathfrak{k}$ which are skew symmetric and satisfy (83) depend only on X and X'. Thus there exists a skew symmetric bilinear map p of $\mathfrak{g}\times\mathfrak{g}$ into $\mathfrak{k}$ such that $r(X,S;\, X',S')=p(X,X')$ for all X, $X' \in \mathfrak{g}$ and S, $S' \in \mathfrak{k}$. We have then

$$[(X,S),(X',S')] = ([X,X'],p(X,X')). \tag{84}$$

If we now use the fact that $[.,.]$ satisfies the Jacobi identities, we see that (78) is satisfied by p. In other words, the extension $\mathfrak{h}$ is equivalent to $\mathfrak{h}_p$.

Suppose now that $\mathfrak{h}_{p_1}$ and $\mathfrak{h}_{p_2}$ are two extensions of $\mathfrak{k}$ by $\mathfrak{g}$ corresponding to the bilinear maps p_1 and p_2. It is easily seen that for any linear map q of $\mathfrak{g}$ into $\mathfrak{k}$, the map

$$L_q : (X,S) \to (X,q(X)+S) \tag{85}$$

is a linear map of $\mathfrak{h}_{p_1}$ into $\mathfrak{h}_{p_2}$, sends $(0,S)$ to $(0,S)$, and is such that the diagram

$$\begin{tikzcd} \mathfrak{h}_{p_1} \arrow[rr, "L_q"] \arrow[dr, "\beta_0"'] & & \mathfrak{h}_{p_2} \arrow[dl, "\beta_0"] \\ & \mathfrak{g} & \end{tikzcd} \tag{86}$$

is commutative; and that every linear map L of $\mathfrak{h}_{p_1}$ into $\mathfrak{h}_{p_2}$, which sends $(0,S)$ to $(0,S)$, and which is such that (86) is commutative (with L

instead of L_q), is of the form L_q for suitable q. It is then clear that L_q is a linear isomorphism of $\mathfrak{h}_{p_1}$ onto $\mathfrak{h}_{p_2}$. A trivial calculation shows that L_q is a homomorphism of the *Lie algebras* $\mathfrak{h}_{p_1}$ and $\mathfrak{h}_{p_2}$ if and only if p_1, p_2 and q are related as in (81). This completes the proof of the lemma.

We shall call a skew symmetric bilinear map p of $\mathfrak{g} \times \mathfrak{g}$ into $\mathfrak{k}$ *closed* if it satisfies (78). We shall say that p is *exact* if there exists a linear map q of $\mathfrak{g}$ into $\mathfrak{k}$ such that

$$p(X,Y) = q([X,Y]) \qquad (X,\ Y \in \mathfrak{g}). \tag{87}$$

The vector space of all closed skew symmetric bilinear maps of $\mathfrak{g} \times \mathfrak{g}$ into $\mathfrak{k}$ is denoted by $\mathfrak{M}_{\mathfrak{k}}'(\mathfrak{g})$, and the subspace of exact elements is denoted by $\mathfrak{E}_{\mathfrak{k}}(\mathfrak{g})$. We write

$$\mathfrak{M}_{\mathfrak{k}}(\mathfrak{g}) = \mathfrak{M}_{\mathfrak{k}}'(\mathfrak{g})/\mathfrak{E}_{\mathfrak{k}}(\mathfrak{g}). \tag{88}$$

Corollary 7.35. *For a* $p \in \mathfrak{M}_{\mathfrak{k}}'(\mathfrak{g})$, *$p$ is exact if and only if* $\{0\} \times \mathfrak{k}$ *is a semi-direct summand of* $\mathfrak{h}_p$.

Proof. If p is defined by (87) for some linear map q of $\mathfrak{g}$ into $\mathfrak{k}$, it is immediate that the range $\mathfrak{b}$ of the map $X \to (X,q(X))$ is a subalgebra of $\mathfrak{h}_p$ such that $\alpha_0[\mathfrak{k}] \cap \mathfrak{b} = 0$ and $\alpha_0[\mathfrak{k}] + \mathfrak{b} = \mathfrak{h}_p$. Conversely, let $\mathfrak{b}$ be such a Lie subalgebra. Then for each $X \in \mathfrak{g}$, let us define the element $q(X) \in \mathfrak{k}$ by the condition that $(X,q(X)) \in \mathfrak{b}$. q is evidently a linear map of $\mathfrak{g}$ into $\mathfrak{k}$, and the condition $[(X,q(X)),(Y,q(Y))] = ([X,Y],q([X,Y]))$ leads to (87).

The theory developed so far has established a canonical isomorphism between the groups $M_K(G)$ and $M_{K^*}(G)$, and also a canonical one-one correspondence between $M_{K^*}(G)$ and $\mathfrak{M}_{\mathfrak{k}}(\mathfrak{g})$ (G simply connected). Now $\mathfrak{M}_{\mathfrak{k}}(\mathfrak{g})$ is a vector space. On the other hand, if we write K^* additively, it is isomorphic to a vector group, so that $M_{K^*}(G)$ also becomes a vector space, with addition as the group operation in it. It is indeed most natural to expect that the canonical correspondence between the elements of $M_{K^*}(G)$ and $\mathfrak{M}_{\mathfrak{k}}(\mathfrak{g})$ is actually a linear isomorphism. We shall now prove that this is indeed so. We assume $K^* = R^n$. The R^n-multipliers for G then satisfy (14) and form a vector space over $\mathbf{R}$, and the exact ones form a subspace of it. $M_{R^n}(G)$ is thus a vector space.

Lemma 7.36. *Let G be connected and simply connected. For any analytic R^n-multiplier m for G, let $[m]$ be the similarity class of m. Let $b_{[m]}$ be the element of $\mathfrak{M}_{R^n}(\mathfrak{g})$ which corresponds to the equivalence class of central extensions of R^n by G which are associated with $[m]$. Then*

$$[m] \to b_{[m]} \tag{89}$$

is a linear isomorphism of $M_{R^n}(G)$ onto $\mathfrak{M}_{R^n}(\mathfrak{g})$.

Proof. Let m be an analytic R^n-multiplier for G and let H_m be the associated group extension. For (x,k) and $(y,k') \in H_m$ $(=G \times_m R^n$ in our earlier notation),

$$(x,k)(y,k') = (xy, k+k'+m(x,y)).$$

Now $(e,0)$ is the identity of H_m, and H_m has the product analytic structure. We shall now determine the Lie algebra $\mathfrak{h}$ of H_m. Following Chevalley [1] we regard elements of $\mathfrak{h}$ as left invariant vector fields on H_m; for each $(x,k) \in H_m$, we identify the tangent space to H_m canonically with $G_x \times R^n$, G_x being the tangent space to G at x. Of course we identify the tangent space to R^n at any of its points with R^n itself in the usual fashion.

Let $X \in \mathfrak{g}$ be the left invariant vector field $x \rightarrow X_x$ on G and let $u \in R^n$. We write (X,u) for the left invariant vector field of H_m which assigns to $(e,0)$ the tangent vector (X_e,u). Since left translation by (x,k) is $(y,k') \rightarrow (xy, k+k'+m(x,y))$, we see that the tangent vector assigned by (X,u) at (x,k) is given by the expression

$$X_x, u + \left(\frac{d}{dt} m(x, \exp tX)\right)_{t=0}.$$

Write now for any C^∞ function f on $G \times G$ and $Y_1, \cdots, Y_r, Z_1, \cdots, Z_s \in \mathfrak{g}$, $y, z \in G$,

$$\begin{aligned} &f(y; Y_1 \cdots Y_r : z; Z_1 \cdots Z_s) \\ &\qquad = (\partial^{r+s}\varphi/\partial t_1 \cdots \partial t_r \partial u_1 \cdots \partial u_s)_{t_1 = \cdots = t_r = u_1 = \cdots = u_s = 0}, \end{aligned} \tag{90}$$

where

$$\begin{aligned} &\varphi(t_1, \cdots, t_r, u_1, \cdots, u_s) \\ &\qquad = f(y \exp t_1 Y_1 \cdots \exp t_r Y_r, z \exp u_1 Z_1 \cdots \exp u_s Z_s). \end{aligned}$$

Then (X,u) is given by

$$(X,u) : x, k \rightarrow (X_x, u + m(x : e; X)). \tag{91}$$

From (91) we find

$$[(X,u),(Y,v)] : x, k \rightarrow ([X,Y]_x, m(x; X : e; Y) - m(x; Y : e; X)). \tag{92}$$

From the identities (14) we obtain, on differentiation,

$$m(x; X : e; Y) = m(x : e; XY) + m(e; X : e; Y).$$

This then leads to

$$\begin{aligned} &[(X,u),(Y,v)] : x, k \rightarrow \\ &\qquad ([X,Y]_x, m(x : e; [X,Y]) + m(e; X : e; Y) - m(e; Y : e; X)) \end{aligned} \tag{93}$$

so that finally, on comparing (93) with (91), we obtain

$$[(X,u),(Y,v)] = ([X,Y], m(e; X : e; Y) - m(e; Y : e; X)). \tag{94}$$

In other words, if we write

$$B_m(X,Y) = m(e;\ X : e;\ Y) - m(e;\ Y : e;\ X) \tag{95}$$

then the Lie algebra $\mathfrak{h}$ of H_m can be identified with $\mathfrak{g} \times R^n$ in a canonical fashion and

$$[(X,u),(Y,v)] = ([X,Y],B_m(X,Y)). \tag{96}$$

B_m is evidently a skew symmetric bilinear map of $\mathfrak{g} \times \mathfrak{g}$ into R^n. The Jacobi identity applied to (96) leads as before to the fact that B_m is a closed form.

We know already that $[m] \to b_{[m]}$ is a one-one correspondence of $M_{R^n}(G)$ onto $\mathfrak{M}_{R^n}(\mathfrak{g})$, and (96) shows that $b_{[m]}$ is the element of $\mathfrak{M}_{R^n}(\mathfrak{g})$ defined by B_m. But a look at (95) shows that $B_m(X,Y)$ depends *linearly* on m, for fixed X, $Y \in \mathfrak{g}$. Consequently, $[m] \to b_{[m]}$ is a linear isomorphism of $M_{R^n}(G)$ onto $\mathfrak{M}_{R^n}(\mathfrak{g})$. This proves the lemma.

Collecting all these results together we obtain the following theorem. We write G for the Lie group we have been concerned with. K is a connected abelian Lie group, K^* a connected, simply connected covering group of K, and ε is a covering homomorphism of K^* onto K.

Theorem 7.37. *Let G be a connected and simply connected Lie group, and let K, K^*, and ε be defined as above. Then any K-multiplier for G is similar to one which is analytic on $G \times G$. For any analytic (respectively C^∞) K^*-multiplier m^*, $\varepsilon \circ m^*$ is an analytic (respectively C^∞) K-multiplier for G, and every analytic (respectively C^∞) K-multiplier can be represented in this form; $m_1{}^*$ and $m_2{}^*$ are similar if and only if $\varepsilon \circ m_1{}^*$ and $\varepsilon \circ m_2{}^*$ are similar. If we associate with any equivalence class of central extensions of K by G the equivalence class of associated extensions of $\mathfrak{k}$ by $\mathfrak{g}$, we obtain a one-one correspondence whose range is the set of all equivalence classes of central extensions of $\mathfrak{k}$ by $\mathfrak{g}$. These latter classes are in canonical one-one correspondence with elements of the space $\mathfrak{M}_{\mathfrak{k}}(\mathfrak{g})$. If we write K^* additively and treat it as a real vector space, $M_{K^*}(G)$ becomes a real vector space and the induced correspondence of $M_{K^*}(G)$ with $\mathfrak{M}_{\mathfrak{k}}(\mathfrak{g})$ is a linear isomorphism. Finally, a K-multiplier G is exact if and only if, in the associated extension of $\mathfrak{k}$ by $\mathfrak{g}$, the image of $\mathfrak{k}$ is a semidirect summand.*

5. EXAMPLES

We shall now discuss a number of examples which illuminate various aspects of the theory described in Sections 1 through 4. We restrict ourselves to the case of (ordinary) multipliers and R^1-multipliers.

Covering Groups: There is no direct connection between simple connectedness and exactness. We shall give, further on, examples of connected but not simply connected Lie groups admitting only exact multipliers, and of simply connected Lie groups whose multiplier groups are not trivial. There are, however, some elementary relations between multipliers and representations of a group and its covering group that may be worthwhile putting down.

Let G be a connected Lie group, G^* a connected and simply connected covering group of G, and δ the covering homomorphism. Let $x^* \to V_{x^*}$ be an *irreducible* unitary representation of G^* in a Hilbert space $\mathscr{H}$. By Schur's lemma, V maps every element in the center of G^* into a scalar multiple of the identity. Since the kernel Z of δ is contained in the center of G^*, it follows at once that $V_{z^*x^*} = \chi(z^*)V_{x^*}$ for all $z^* \in Z$ and $x^* \in G^*$, χ being a character of Z. This implies that for any $g \in G$, all the operators V_{g^*} with $\delta(g^*) = g$, define the same element of the projective group of $\mathscr{H}$. In this way, each irreducible representation of G^* induces a projective representation of G.

The relation between G and G^* is also useful in another way. Let m be a multiplier for G and let m^* be the multiplier $x^*, y^* \to m(\delta(x^*), \delta(y^*))$. It is then easily seen that $m \to m^*$ induces a homomorphism, say h, of $M(G)$ into $M(G^*)$. Since G^* is simply connected, theorem 7.37 can be applied to the study of $M(G^*)$. In this way useful global insights into the structure of $M(G)$ may be obtained. However, it must be remembered that h is in general neither injective nor surjective, so that, in this approach to the analysis of $M(G)$ in terms of $M(G^*)$, the kernel and the range of h must both be determined. For example, if G is a torus, $M(G)$ is trivial while $M(G^*)$ is nontrivial as soon as $\dim G^* > 1$; if G is semisimple, $M(G^*)$ is trivial, but $M(G)$ will not be (cf. supra).

$G = \mathbf{R}^s$, *a vector group in s-dimensions.* In this case we can explicitly describe the multiplier group (cf. also Weyl [1], pp. 273–274). We formulate our conclusions in the form of the following theorem. It is an easy consequence of the results described in theorem 7.37.

Theorem 7.38. *For any real skew symmetric bilinear form p on $\mathbf{R}^s \times \mathbf{R}^s$,*

$$m_p : x, y \to \exp ip(x,y) \tag{97}$$

is a multiplier for $\mathbf{R}^s$. The mapping which assigns to p, the similarity class of m_p, is an isomorphism of the additive group of all skew symmetric bilinear forms on $\mathbf{R}^s \times \mathbf{R}^s$ onto the multiplier group $M(\mathbf{R}^s)$. In particular, all multipliers of $\mathbf{R}^1$ are exact.

Proof. It is trivial to check that m_p is a multiplier for any bilinear form p on $\mathbf{R}^s \times \mathbf{R}^s$ (skew symmetric or not) and that $m_{p_1+p_2} = m_{p_1} m_{p_2}$.

To prove the remainder of the theorem we pass to the covering group $\mathbf{R}^1$ of T and use theorem 7.37. Now, for any bilinear form $p, x, y \rightarrow p(x,y)$ is an $\mathbf{R}^1$-multiplier for $\mathbf{R}^s$ and it is exact if and only if it is symmetric (corollary 7.14). In particular, $p_1 \simeq p_2$ if and only if $p_1 - p_2$ is symmetric. If p_1 and p_2 are skew symmetric, it is now obvious that $p_1 \simeq p_2$, if and only if $p_1 = p_2$. If p is any bilinear form, we can always write it as $p_1 + c$, where p_1 is skew symmetric and c is symmetric; then $p \simeq p_1$. These conclusions then carry over to the T-multipliers m_p.

In order to complete the proof we go back to (95). We write $B_p = B_{m_p}$ and prove that the mapping $p \rightarrow B_p$ has the set of *all* skew symmetric bilinear maps of $\mathbf{R}^s \times \mathbf{R}^s$ into $\mathbf{R}^1$ for its range (we identify $\mathfrak{g}$ with $\mathbf{R}^s$ also in the obvious fashion so that exp is the identity map). This is trivial, for, by (95),

$$-iB_p(X, Y) = 2p(X, Y) \qquad (X,\ Y \in \mathbf{R}^s),$$

whenever p is a skew symmetric bilinear form on $\mathbf{R}^s \times \mathbf{R}^s$. The last assertion follows trivially from the fact that 0 is the only skew symmetric bilinear form on $\mathbf{R}^1 \times \mathbf{R}^1$.

Connected Semisimple Lie Groups: Suppose $\mathfrak{g}$ is a semisimple Lie algebra over the reals and $(\mathfrak{h}, \alpha, \beta)$ a central extension by it of the abelian Lie algebra $\mathfrak{k}$. Since $\mathfrak{h}/\alpha[\mathfrak{k}]$ is semisimple, it follows at once that $\alpha[\mathfrak{k}]$ is the radical of $\mathfrak{h}$. By the Levy-Malčev theorem we conclude that there exists a subalgebra $\mathfrak{b}$ of $\mathfrak{h}$ such that $\alpha[\mathfrak{k}] + \mathfrak{b} = \mathfrak{h}$ and $\alpha[\mathfrak{k}] \cap \mathfrak{b} = 0$. This observation also leads at once to the conclusion (cf. theorem 7.37) that if G is a connected and simply connected semisimple Lie group, then every multiplier for G is exact. Since the classical groups are semisimple, this result applies to the universal covering groups of these groups. For example, we may take G_0 to be the connected component of the group of all invertible linear transformations of an s-dimensional vector space leaving invariant a nonsingular symmetric bilinear form. When $s > 2$, G_0 is known to be a semisimple Lie group. If G denotes a universal covering group of G_0, then all multipliers for G are exact. The examples of greatest interest from the physical point of view arise when $s = 3$ and the bilinear form has a positive definite quadratic form, and when $s = 4$ and the bilinear form has a Minkowskian quadratic form. In other words, the universal covering groups of the rotation and Lorentz groups (connected) admit only exact multipliers.

Inhomogeneous Groups: The symmetry groups of space time cannot be semisimple because they contain the space-time translations as a normal subgroup. Therefore, the results derived in the previous example are inadequate for physical applications and it is necessary to extend their scope. We shall give here one such extension.

Let G be a connected Lie group and let $\rho(x \to \rho(x))$ be a (not necessarily unitary) representation of G in a finite dimensional real vector space V. The map $x, v \to \rho(x)v$ is analytic from $G \times V$ into V and converts V into a G-space. The semidirect product

$$G_\rho = G \times_\rho V \tag{98}$$

can then be formed. G_ρ (in the product analytic structure) is a connected Lie group; if G is simply connected, it is a simply connected Lie group. We identify V and G with closed subgroups of G_ρ. V is normal and for $x \in G$, $v \in V$, $xvx^{-1} = \rho(x)v$. We shall call G_ρ the *inhomogeneous group associated with G and ρ.*

Our purpose now is to prove a theorem from which it follows that under certain conditions the covering group of the inhomogeneous group admits only exact multipliers. We shall call the representation ρ of G in a real vector space V *admissible* if no nonzero skew symmetric bilinear form on $V \times V$ is invariant under ρ.

The fact that many representations of a group are admissible is contained in the following lemma. Since we work with real vector spaces it has to be formulated in the context of absolutely irreducible representations. A representation $\pi(x \to \pi(x))$ of G in a real vector space W is said to be *absolutely irreducible* if the only endomorphisms of W which commute with all the $\pi(x)$ are scalar multiples of the identity. If G is semisimple, absolute irreducibility implies irreducibility in the usual geometric sense and absolute irreducibility of π is equivalent to the irreducibility of π in the complexification of W (G semisimple). The classical Schur's lemma can be formulated for absolutely irreducible representations as follows: if π and π' are absolutely irreducible representations of a semisimple group G, acting in W and W', then the vector space of linear maps L from W into W' such that $L\pi(x) = \pi'(x)L$ for all $x \in G$ is of dimension ≤ 1, the dimension being 1 if and only if π and π' are equivalent.

Lemma 7.39. *Let ρ be a representation of a connected semisimple Lie group G in V. Suppose that V is a direct sum of ρ-invariant subspaces $V_1, \cdots, V_s$ such that for each i,* (i) *the subrepresentation ρ_i (of ρ) defined by V_i is absolutely irreducible and* (ii) *there exists a nonzero symmetric bilinear form on $V_i \times V_i$ which is invariant under ρ_i. If the representations $\rho_1, \rho_2, \cdots, \rho_s$ are mutually inequivalent, then ρ is admissible.*

Proof. Let V^* be the dual of V and for $x \in G$ and $t \in V^*$, let $\rho^*(x)t$ be the element $v \to t(\rho(x^{-1})v)$ of V^*. Then $\rho^*(x)(t \to \rho^*(x)t)$ is easily verified to be an invertible endomorphism of V^*. It is also easy to verify that $\rho^* : x \to \rho^*(x)$ is a representation of G in V^*, the so-called contragredient representation acting in V^*. Now any bilinear form B on $V \times V$ gives rise to a linear transformation T_B of V into V^*, and B is invariant if and only if T_B *intertwines* ρ and ρ^*, i.e.,

$$T_B\rho(x) = \rho^*(x)\, T_B \tag{99}$$

for all $x \in G$. Let $V_i^* \subseteq V^*$ be the annihilator of $\sum_{j \neq i} V_j$. Then ρ^* leaves each V_i^* invariant. From the fact that ρ_i leaves a nonzero bilinear form on $V_i \times V_i$ invariant, it follows that there exists a nonzero linear map T_i, mapping V_i into V_i^* and V_j into 0 for $j \neq i$, such that T_i intertwines ρ and ρ^*. From the standard form of Schur's lemma we conclude, using the mutually inequivalent nature of $\rho_1, \cdots, \rho_s$, that the operators $T_1, \cdots, T_s$ span the vector space of operators which intertwine ρ and ρ^*. This implies that every operator which intertwines ρ and ρ^* is a T_B for a *symmetric* bilinear form B on $V \times V$. Hence any bilinear form on $V \times V$ left invariant by ρ is necessarily symmetric. Consequently, ρ is admissible.

Theorem 7.40. *Let G be a connected semisimple Lie group and ρ an admissible representation of G in a real finite dimensional vector space. Then, for the universal covering group G_ρ^* of G_ρ, every multiplier is exact. In particular, if $\alpha(x \to \alpha_x)$ is a representation of G_ρ into the group of automorphisms of the logic of a (complex separable) Hilbert space $\mathscr{H}$, there exists a representation $U(x^* \to U_{x^*})$ of G^* in $\mathscr{H}$ such that for each x in G, α_x is induced by U_{x^*} for any x^* in G_ρ^* lying above x. If ρ does not contain the trivial representation as a subrepresentation, U is the only representation of G_ρ^* with this property.*

Proof. Let us first describe G_ρ^* in a specific manner. Let G^* be a connected and simply connected covering group of G and let $\varepsilon(x^* \to x)$ be the covering homomorphism. Then $\rho \circ \varepsilon = \rho^* : x^* \to \rho(\varepsilon(x^*))$ is a representation of G^* in V. ρ^* is obviously admissible and we form the group

$$G_\rho^* = G^* \times_{\rho^*} V. \tag{100}$$

It is clear that G_ρ^* is connected and simply connected and the map $x^*, v \to \varepsilon(x^*), v$ is the covering map of G_ρ^* on G_ρ (we identify V and G^* with subgroups of G_ρ^* as usual). If α is the given representation of G_ρ in the group of automorphisms of the logic of $\mathscr{H}$, we know (cf. footnote in Section 1) that there is a projective representation V^1 of G_ρ in $\mathscr{H}$ which induces α. We lift V^1 to the projective representation $\tilde{V} : x^*, v \to V^1_{\varepsilon(x^*),v}$ of G_ρ^*. We now apply theorem 7.5. In order that α be induced by a representation of G_ρ^* in the manner described in the theorem, it is necessary and sufficient that there is a unitary representation of G_ρ^* which gives rise to the same representation as $\tilde{V}$ in the projective group of $\mathscr{H}$; clearly this will happen if and only if the multiplier of $\tilde{V}$ is exact. Also, the representation of G_ρ^* which induces α is unique up to multiplication by a continuous homomorphism of $G_\rho{}^*$ into the group T (corollary 7.7). Thus, in order to prove the assertions of the theorem, it is enough to prove two things: first, that all multipliers for $G_\rho{}^*$ are exact, and second, that $G_\rho{}^*$ admits no nontrivial continuous homomorphisms into T under the assumptions made on ρ.

We take up the first point. Let $\mathfrak{g}$, $\mathfrak{v}$, and $\mathfrak{g}_\rho$ be the Lie algebras of G^*, V, and $G_\rho{}^*$, respectively. $\mathfrak{v}$ is an ideal in $\mathfrak{g}_\rho$, $\mathfrak{g}$ a subalgebra in $\mathfrak{g}_\rho$, and

$$\begin{aligned} \mathfrak{v} \cap \mathfrak{g} &= 0, \\ \mathfrak{v}+\mathfrak{g} &= \mathfrak{g}_\rho. \end{aligned}$$

If $\dot\rho$ is the representation of $\mathfrak{g}$ in V that corresponds to the representation ρ^* of G^*, we shall, by using the canonical isomorphism of $\mathfrak{v}$ on V, assume that ρ and $\dot\rho$ act also on $\mathfrak{v}$. Thus, for $Z \in \mathfrak{v}$ and $X \in \mathfrak{g}$,

$$[X,Z] = \dot\rho(X)(Z); \tag{101}$$

this corresponds to the relation $x^*vx^{*-1} = \rho(\varepsilon(x^*))v$ ($x^* \in G^*$, $v \in V$).

In order to prove that all multipliers for $G_\rho{}^*$ are exact, it is sufficient to prove that if $\mathfrak{k}$ is any abelian Lie algebra and $(\mathfrak{h},p,q)$ a central extension of $\mathfrak{k}$ by $\mathfrak{g}_\rho$, then $p[\mathfrak{k}]$ is a semidirect summand of $\mathfrak{h}$. We shall now demonstrate this. We write

$$\mathfrak{v}' = q^{-1}(\mathfrak{v}) \tag{102}$$

and

$$\mathfrak{g}' = q^{-1}(\mathfrak{g}). \tag{103}$$

q maps $\mathfrak{g}'$ onto $\mathfrak{g}$ and the kernel of q is $p[\mathfrak{k}]$. Hence we conclude from lemma 7.25 that there exists a subalgebra $\mathfrak{a}$ such that $\mathfrak{a} \cap p[\mathfrak{k}] = 0$ and $\mathfrak{a} + p[\mathfrak{k}] = \mathfrak{g}'$. q is an isomorphism of $\mathfrak{a}$ onto $\mathfrak{g}$, so that $\mathfrak{a}$ is semisimple. Since $\mathfrak{v} \cap \mathfrak{g} = 0$, $\mathfrak{v}' \cap \mathfrak{a} \subseteq p[\mathfrak{k}]$ so that $\mathfrak{v}' \cap \mathfrak{a} = 0$. In other words,

$$\begin{aligned} \mathfrak{v}' \cap \mathfrak{a} &= 0, \\ \mathfrak{v}'+\mathfrak{a} &= \mathfrak{h}. \end{aligned} \tag{104}$$

Now $\mathfrak{v}'$ is an ideal in $\mathfrak{h}$, and so for any $X \in \mathfrak{a}$, $(\operatorname{ad} X)(Z) = [X,Z]$ lies in $\mathfrak{v}'$ whenever $Z \in \mathfrak{v}'$. Thus $X \to \operatorname{ad} X$ defines a representation of $\mathfrak{a}$ in $\mathfrak{v}'$. As $p[\mathfrak{k}]$ is central, ad X annihilates all the elements of $p[\mathfrak{k}]$. Therefore, as $\mathfrak{a}$ is semisimple, there exists, by Weyl's theorem (cf. Chevalley [3], pp. 70–73) a subspace $\mathfrak{b}_1$ such that $\mathfrak{b}_1 \cap p[\mathfrak{k}] = 0$, $\mathfrak{b}_1 + p[\mathfrak{k}] = \mathfrak{v}'$, and $\mathfrak{b}_1$ is invariant under all ad X, $X \in \mathfrak{a}$. Let

$$\mathfrak{b} = \mathfrak{b}_1 + \mathfrak{a}. \tag{105}$$

$\mathfrak{b}$ is a linear manifold, $\mathfrak{b} \cap p[\mathfrak{k}] = 0$, and $\mathfrak{b} + p[\mathfrak{k}] = \mathfrak{h}$. Our claim that $p[\mathfrak{k}]$ is a semidirect summand will be proved if we show that $\mathfrak{b}$ is a *subalgebra* of $\mathfrak{h}$. Since $[X,Y] \in \mathfrak{b}_1$ whenever $X \in \mathfrak{a}$ and $Y \in \mathfrak{b}_1$, and since $\mathfrak{a}$ is a subalgebra, we have only to show that $[X,Y] \in \mathfrak{b}_1$ whenever X, Y are in $\mathfrak{b}_1$. We shall actually prove that $[X,Y]=0$ for X, Y in $\mathfrak{b}_1$. Now q maps $\mathfrak{b}_1$ onto $\mathfrak{v}$, so that the abelian nature of $\mathfrak{v}$ implies that, for $X, Y \in \mathfrak{b}_1$, $[X,Y] \in p[\mathfrak{k}]$. We write

$$K_1(X,Y) = p^{-1}([X,Y]) \qquad (X,\ Y \in \mathfrak{b}_1). \tag{106}$$

K_1 is evidently a skew symmetric bilinear map of $\mathfrak{b}_1 \times \mathfrak{b}_1$ into $\mathfrak{k}$. If $Z \in \mathfrak{a}$ and $X, Y \in \mathfrak{b}_1$, $[X,[Y,Z]]+[Y,[Z,X]]=0$, since this element is nothing but $-[Z,[X,Y]]$, and hence 0 as $[X,Y]$ belongs to the center of $\mathfrak{h}$. Therefore,

$$K_1([Z,Y],X)+K_1(Y,[Z,X]) = 0. \tag{107}$$

Now q is an isomorphism of the Lie algebra $\mathfrak{a}$ onto $\mathfrak{g}$, and the vector space $\mathfrak{b}_1$ onto $\mathfrak{v}$. Therefore there exists a unique skew symmetric bilinear map K of $\mathfrak{v} \times \mathfrak{v}$ into $\mathfrak{k}$ such that

$$K_1(X',Y') = K(q(X'),q(Y')) \qquad (X',\ Y' \in \mathfrak{b}_1) \tag{108}$$

and

$$K([Z,Y],X)+K(Y,[Z,X]) = 0 \tag{109}$$

for all X, Y in $\mathfrak{v}$ and Z in $\mathfrak{g}$. In other words,

$$K(\dot{\rho}(Z)Y,X)+K(Y,\dot{\rho}(Z)X) = 0. \tag{110}$$

Since (110) is valid for all Z in $\mathfrak{g}$, we conclude that K is invariant under the representation ρ. If we choose a basis $\{A_1,\cdots, A_n\}$ for $\mathfrak{k}$ and write

$$K(X,Y) = \sum_{i=1}^{n} K_i(X,Y)A_i,$$

then each K_i is an invariant skew symmetric bilinear form on $\mathfrak{v} \times \mathfrak{v}$. Each K_i is 0, as ρ is admissible. This implies that $K=0$. Therefore, $K_1=0$. This proves that $\mathfrak{b}=\mathfrak{b}_1+\mathfrak{a}$ is a subalgebra of $\mathfrak{h}$. In other words, $p[\mathfrak{k}]$ is a semidirect summand of $\mathfrak{h}$.

We finally come to the second point, namely, to the proof that under the conditions mentioned in the theorem $G_\rho{}^*$ admits no one dimensional representation other than the trivial one. If it does, $\mathfrak{g}_\rho=\mathfrak{v}+\mathfrak{g}$ will admit a nontrivial one-dimensional representation. Therefore there will exist a (complex valued) nonzero linear function λ on $\mathfrak{g}_\rho$ such that $\lambda([X,Y])=0$ for all X, Y in $\mathfrak{g}_\rho$. Since $\mathfrak{g}$ is semisimple, every element of $\mathfrak{g}$ is a linear combination of elements of the form $[X,Y]$ with X, Y in $\mathfrak{g}$ (cf. Chevalley [3], p. 67) so that $\lambda=0$ on $\mathfrak{g}$. Let λ' be the restriction of λ to $\mathfrak{v}$. Then the equation $\lambda([Z,X])=0$, for $X \in \mathfrak{v}$ and $Z \in \mathfrak{g}$, shows that the null space of λ' is invariant under $\dot{\rho}$. Since $\dot{\rho}$ is semisimple, we can select a subspace complementary to the null space of λ' and left invariant by $\dot{\rho}$. This space is one-dimensional and so, as $\mathfrak{g}$ is semisimple, we obtain a nonzero vector fixed by $\dot{\rho}$. This is a contradiction. The proof of the theorem is now complete.

This is a basic theorem insofar as relativistic applications are concerned. It tells us, for example, that any representation of the (connected) inhomogeneous Lorentz group can be obtained from a unitary representation of the covering group. To see this, all we have to do is take V as R^4, G as the connected component of the group of Lorentz matrices acting on V, and $x \rightarrow \rho(x)$ as the action of G in V. It is quite easy to check that ρ is admissible. In this special case, theorem 7.40 was obtained by Wigner [1] and, in his work, this theorem was the point of departure for his classification of relativistic wave equations. In the special case when G is the group of *all* invertible linear transformations of V leaving a nondegenerate quadratic form on V invariant ($\dim V > 2$), theorem 7.40 was obtained by Bargmann [1].

Compact Groups. Much is known about multipliers for compact groups (cf. Moore [1]). We shall content ourselves with the proof of a simple result.

Theorem 7.41. *Let G be any compact group satisfying the second axiom of countability. Then its multiplier group is a torsion group, i.e., every element of $M(G)$ is of finite order. If G is a torus, or a connected and simply connected compact Lie group, G admits only exact multipliers.*

Proof. Let m be a multiplier for G. By corollary 7.18 there exists a finite dimensional m-representation $\rho(x \rightarrow \rho(x))$, acting in V say. Let $\zeta \rightarrow \zeta^{1/n}$ denote a Borel nth root of ζ, i.e., a Borel map of T into itself which fixes 1 and whose nth power is the coordinate function $\zeta \rightarrow \zeta$, n being the dimension of V. Then

$$\rho' : x \rightarrow (\det \rho(x))^{1/n} \rho(x)$$

is a projective representation of G; and ρ' defines the same homomorphism of G into the projective group of V, as ρ. If m' is the multiplier of ρ', then m and m' are similar. But $\det \rho'(x) = 1$ for all x so that for all x, y in G, $[m'(x,y)]^n = 1$. Hence $(m)^n$ is exact, proving that any element of $M(G)$ has finite order.

If G is a connected and simply connected compact Lie group, then G is semisimple by a theorem of H. Weyl (cf. Helgason [1]), so that G admits only exact multipliers.

Suppose finally that G is a torus, say $G = T^r$. Let $\varepsilon : v \rightarrow \varepsilon(v)$ be the covering homomorphism from $\mathbf{R}^r$ onto T^r. If m is any multiplier for G, $m_\varepsilon : v, v' \rightarrow m(\varepsilon(v), \varepsilon(v'))$ is a multiplier for $\mathbf{R}^r$. Since m^s is exact from some integer $s \geq 1$, so is $m_\varepsilon{}^s$. But the multiplier group of $\mathbf{R}^r$ is canonically isomorphic to the $\mathbf{R}^1$-multiplier group of $\mathbf{R}^r$ which is a vector group. Hence $m_\varepsilon{}^s$ cannot be exact unless m_ε is itself exact. Therefore, there is a Borel function a from $\mathbf{R}^r$ to T such that $m_\varepsilon(v,v') = a(v+v')/a(v)a(v')$. Since ε is a local group isomorphism, we can find a Borel function b on T^r, with values in T, and an open set N of T^r containing the identity, such that

$$m(x,y) = \frac{b(xy)}{b(x)b(y)} \tag{111}$$

for all $x, y \in N \times N$. Equation (111) implies that m is exact (cf. corollary 7.14).

Groups Which Are Not Semisimple. If the Lie group G is not semisimple, then it may admit nonexact multipliers. We shall now sketch the construction of the multiplier group for the inhomogeneous Galilean group. We closely follow Bargmann [1] in this discussion.

If (x,t) $(x = x_1, x_2, x_3)$ denotes the coordinates of space-time points, then an inhomogeneous Galilean transformation is a map

$$r : x, t \to x', t', \tag{112}$$

where

$$\begin{aligned} x' &= Wx + tv + u, \\ t' &= t + \eta, \end{aligned} \tag{113}$$

W being a rotation in 3-space ($\det W = +1$), u, v vectors in 3-space, and η a real number. The set of all such transformations forms, under composition, the so-called inhomogeneous Galilean group; we have, for the multiplication in this group,

$$(W,\eta,v,u)(W',\eta',v',u') = (WW',\eta+\eta',Wv'+v,Wu'+u+\eta'v). \tag{114}$$

If we associate with r the matrix

$$M(r) = \begin{pmatrix} W & v & u \\ 0 & 1 & \eta \\ 0 & 0 & 1 \end{pmatrix}, \tag{115}$$

then $r \to M(r)$ is a faithful representation. The Galilean group is a semidirect product; the subgroup A of all translations in space-time,

$$A = \{(1,\eta,0,u)\} \tag{116}$$

is normal, the subgroup G_0 of transformations

$$G_0 = \{(W,0,v,0)\} \tag{117}$$

is a closed subgroup; A and G_0 generate the whole inhomogeneous group. In fact,

$$(1,\eta,0,u)(W,0,v,0) = (W,\eta,v,u) \tag{118}$$

and

$$(W,0,v,0)(1,\eta,0,u)(W,0,v,0)^{-1} = (1,\eta,0,Wu+\eta v), \tag{119}$$

which exhibits explicitly the semidirect product nature. We denote the inhomogeneous group by G. The subgroup G_0, unlike the Lorentz group, is not semisimple, and this accounts for some of the more subtle differences between the Galilean theory and Lorentzian theory. Let M and H denote the subgroups of G_0 given by

$$M = \{(1,0,v,0)\}, \tag{120}$$

$$H = \{(W,0,0,0)\}. \tag{121}$$

M is a closed normal subgroup of G_0, H is a subgroup with $G_0 = MH$; moreover,

$$(W,0,0,0)(1,0,v,0)(W,0,0,0)^{-1} = (1,0,Wv,0), \tag{122}$$

so that G_0 is once again a semidirect product, with respect to the action of H in M given by (122).

Note also that the space motions of the form $(W,0,0,u)$ constitute the subgroup $S = UH$, where U is the subgroup (of all translations in space) $\{(1,0,0,u)\}$. U is a normal subgroup of S and

$$(W,0,0,0)(1,0,0,u)(W,0,0,0)^{-1} = (1,0,0,Wu); \tag{123}$$

therefore once again we have a semidirect product.

The first step in determining the multiplier group of G is to examine the central extensions at the Lie algebra level. Theorem 7.37 reduces this to the question of determining the closed skew symmetric bilinear forms on $\mathfrak{g} \times \mathfrak{g}$, modulo the exact ones. We shall now do this. As is customary, we denote the subalgebras of $\mathfrak{g}$ by the German letters which correspond to the Latin letters that denote the analytic subgroups of G. The relations (116) through (123) give considerable insight into the structure of $\mathfrak{g}$. $\mathfrak{a}$ is an ideal in $\mathfrak{g}$, $\mathfrak{g} = \mathfrak{a} + \mathfrak{g}_0$, $\mathfrak{a} \cap \mathfrak{g}_0 = 0$; $\mathfrak{m}$ is an ideal in $\mathfrak{g}_0$, $\mathfrak{g}_0 = \mathfrak{m} + \mathfrak{h}$, $\mathfrak{m} \cap \mathfrak{h} = 0$; $\mathfrak{g}_0$ and $\mathfrak{h}$ are subalgebras of $\mathfrak{g}$. Also $\mathfrak{s} = \mathfrak{u} + \mathfrak{h}$ is a subalgebra, $\mathfrak{u}$ is an ideal in $\mathfrak{s}$, $\mathfrak{u} \cap \mathfrak{h} = 0$.

Since A, M, and U are abelian and since H acts on all of them via the inner automorphisms, H acts naturally on $\mathfrak{a}$, $\mathfrak{m}$, and $\mathfrak{u}$. The differential of this action is the usual adjoint action of $\mathfrak{h}$ on $\mathfrak{a}$, $\mathfrak{m}$, $\mathfrak{u}$. Thus, using the standard notation according to which $[\mathfrak{p},\mathfrak{p}']$ is the linear space spanned by all elements of the form $[X,Y]$ with $X \in \mathfrak{p}$ and $Y \in \mathfrak{p}'$,

$$[\mathfrak{h},\mathfrak{m}] \subseteq \mathfrak{m}, \qquad [\mathfrak{h},\mathfrak{u}] \subseteq \mathfrak{u}, \tag{124}$$

and

$$[\mathfrak{h},\mathfrak{p}] = 0, \tag{125}$$

where $\mathfrak{p}$ is the one-dimensional subalgebra of $\mathfrak{g}$ which corresponds to the subgroup $\{(1,\eta,0,0)\}$ of A.

Moreover, from (122) and (123) we see that the representations of H in $\mathfrak{m}$ and $\mathfrak{u}$ are equivalent and absolutely irreducible and moreover, leave the Euclidean inner product on these invariant; these are therefore admissible by lemma 7.39. By a direct calculation or by use of the irreducibility of these representations of $\mathfrak{h}$, we find

$$[\mathfrak{h},\mathfrak{m}] = \mathfrak{m}, \qquad [\mathfrak{h},\mathfrak{u}] = \mathfrak{u}. \tag{126}$$

Moreover an easy computation shows that

$$[\mathfrak{p},\mathfrak{m}] \subseteq \mathfrak{u}, \qquad [\mathfrak{u},\mathfrak{m}] = 0. \tag{127}$$

Let B be any closed skew symmetric bilinear form on $\mathfrak{g}\times\mathfrak{g}$. Then, for X, $Y \in \mathfrak{a}$ and $Z \in \mathfrak{g}_0$, we have, as $[X,Y]=0$,

$$B([Z,X],Y)+B(X,[Z,Y]) = 0. \tag{128}$$

Equation (128) shows that B is invariant under the representation of G_0 in $\mathfrak{a}$. Since the action of G_0 on A leaves no nonzero bilinear form on $A\times A$ invariant, (128) implies at once that $B=0$ on $\mathfrak{a}\times\mathfrak{a}$. Note that this has been proved for *any* closed form B. Next, we observe that $\mathfrak{g}_0=\mathfrak{m}+\mathfrak{h}$ satisfies the conditions of theorem 7.40. In fact, the group H is semisimple and its action on M (which is isomorphic to 3-space) as given by (122) is, as observed by us already, admissible. Hence the restriction of B to $\mathfrak{g}_0\times\mathfrak{g}_0$ is exact. Replacing B by a form which differs from B by an exact one, we may thus assume that

$$B = 0 \quad \text{on} \quad \mathfrak{a}\times\mathfrak{a} \quad \text{and} \quad \mathfrak{g}_0\times\mathfrak{g}_0. \tag{129}$$

We shall study the form B on $\mathfrak{a}\times\mathfrak{g}_0$. Now $\mathfrak{a}+\mathfrak{h}$ is a Lie algebra and $\mathfrak{a}$ is an ideal in it. Moreover, the action of H on A, as given by (119), is admissible by lemma 7.39, since A decomposes into a direct sum of U and the one-dimensional space of time translations. Hence theorem 7.40 applies once again and we conclude that the restriction of B to $(\mathfrak{a}+\mathfrak{h})\times(\mathfrak{a}+\mathfrak{h})$ is exact. Therefore, there exists a linear function q on $\mathfrak{g}$, vanishing on $\mathfrak{m}$, such that $B(X,Y)=q([X,Y])$ for all X, Y in $\mathfrak{a}+\mathfrak{h}$. Since $\mathfrak{h}$ is semisimple, $[\mathfrak{h},\mathfrak{h}]=\mathfrak{h}$ so that (129) implies that q vanishes also on $\mathfrak{h}$. If we define B' by

$$B'(X,Y) = B(X,Y)-q([X,Y]) \qquad (X,\ Y \in \mathfrak{g}\times\mathfrak{g}), \tag{130}$$

then $B-B'$ is exact and

$$B' = 0 \quad \text{on} \quad \mathfrak{a}\times\mathfrak{a},\ \mathfrak{a}\times\mathfrak{h}, \quad \text{and} \quad \mathfrak{g}_0\times\mathfrak{g}_0. \tag{131}$$

If $X \in \mathfrak{p}$, $Y \in \mathfrak{m}$ and $Z \in \mathfrak{h}$, the identity

$$B'(X,[Z,Y])+B'(Z,[Y,X])+B'(Y,[X,Z]) = 0$$

implies, in view of (131) and the relations $[Y,X] \in \mathfrak{u}$, $[X,Z]=0$ (cf. (127)), that

$$B'(X,[Z,Y]) = 0. \tag{132}$$

From (126) and (132) we conclude that

$$B' = 0 \quad \text{on} \quad \mathfrak{p}\times\mathfrak{m}. \tag{133}$$

We are now ready to prove that the set of closed forms, modulo the set of exact forms, is *one* dimensional. To this end, we must analyze the restriction of B' to $\mathfrak{u}\times\mathfrak{m}$. Now, for $Z \in \mathfrak{h}$, we have, as $[\mathfrak{u},\mathfrak{m}]=0$,

$$B'([Z,X],Y)+B'(X,[Z,Y]) = 0 \tag{134}$$

for all X, $Y \in \mathfrak{u}\times\mathfrak{m}$. Since H induces *equivalent* representations on $\mathfrak{u}$ and $\mathfrak{m}$, the vector space of bilinear forms on $\mathfrak{u}\times\mathfrak{m}$ has a one-dimensional subspace of bilinear forms invariant under the natural action of H. Let C_0 be a basis for this space. Let C be defined as the unique symmetric bilinear form on $\mathfrak{g}\times\mathfrak{g}$ such that $C=C_0$ on $\mathfrak{u}\times\mathfrak{m}$ and $C=0$ on all of $\mathfrak{a}\times\mathfrak{a}$, $\mathfrak{g}_0\times\mathfrak{g}_0$, $\mathfrak{a}\times\mathfrak{h}$, and $\mathfrak{p}\times\mathfrak{m}$. Then (131), (133), and (134) show that B' is a multiple of C. In other words, the vector space of closed forms modulo the subspace of exact forms has dimension ≤ 1. On the other hand, it is easily verified that C is actually closed. If C were exact, and $C(X,Y)=q([X,Y])$ for all X and Y in $\mathfrak{g}$, then the equation $[\mathfrak{u},\mathfrak{m}]=0$ would imply that $C=0$ on $\mathfrak{u}\times\mathfrak{m}$, a contradiction. Thus the space $\mathfrak{M}_{R^1}(\mathfrak{g})$ has dimension exactly 1.

Therefore there exists a nonexact $\mathbf{R}^1$-multiplier $m_0{}^*$ for the covering group G^* of G such that every multiplier for G^* is similar to one of the form

$$x^*, y^* \to \exp[i\tau m_0{}^*(x^*,y^*)]$$

for a suitable real number τ.

Actually $m_0{}^*$ can be chosen to induce a nonexact $\mathbf{R}^1$-multiplier m_0 for G itself. Bargmann [1] has indicated a simple method for constructing m_0. Let us write $\xi=(W,\eta)$, $\zeta=(v,u)$. Then we can expand m_0 as a series $\sum_n \varphi_n$, where φ_n is a polynomial in ζ_1, ζ_2, homogeneous of degree n and having coefficients which are functions of ξ_1 and ξ_2. Since the ζ's are transformed linearly in the group multiplication, it follows that each φ_n satisfies the multiplier identities *locally*. This tells us that a simple and promising class of functions, in which we may search for the multipliers, would be the class of functions which are analytic in ξ_1, ξ_2, ζ_1, ζ_2, and which are *bilinear* in ζ_1 and ζ_2 for fixed ξ_1, ξ_2. If we assume that m_0 is of this form and write down the multiplier identities, we obtain equations which are quite easy to solve. We find

$$m_0(r,r') = \langle u,Wv'\rangle - \langle v,Wu'\rangle + \eta'\langle v,Wv'\rangle, \tag{135}$$

where $\langle .,.\rangle$ denotes the Euclidean inner product on 3-space and

$$r = (W,\eta,v,u), \qquad r' = (W',\eta',v',u'). \tag{136}$$

It is of course necessary to check that m_0, when lifted to a $\mathbf{R}^1$-multiplier $m_0{}^*$ for G^*, is not exact. But suppose $m_0{}^*$ is exact. Then $m_0{}^*(x^*,y^*)=m_0{}^*(y^*,x^*)$ whenever x^* and y^* commute, and hence $m_0(x,y)=m_0(y,x)$ if x and y commute and are close to the identity. We contradict this by choosing $x_s=(1,0,sv,0)$ and $y_s=(1,0,0,su)$, with s near 0; $x_s y_s = y_s x_s$ but $m_0(x_s,y_s)-m_0(y_s,x_s)=-2s^2\langle u,v\rangle$.

Theorem 7.42. *Let G be the inhomogeneous Galilean group. Then for any real number* τ

$$m_\tau : r, r' \to \exp i\frac{\tau}{2}[\langle u,Wv'\rangle - \langle v,Wu'\rangle + \eta'\langle v,Wv'\rangle] \tag{137}$$

is a multiplier for G, where

$$r = (W,\eta,v,u) \quad \text{and} \quad r' = (W',\eta',v',u').$$

Every multiplier of G is similar to one of this form. The m_τ, *for distinct* τ, *are not similar. In particular,* m_τ *is nonexact for* $\tau \neq 0$.

NOTES ON CHAPTER VII

For a closely related treatment of multipliers see K. R. Parthasarathy, *Multipliers on locally compact groups*, Lecture Notes in Mathematics, No. 93, Springer-Verlag, Berlin, 1969. For a systematic treatment of the cohomology of locally compact second countable groups see the papers of C. C. Moore: Trans. Amer. Math. Soc., **113** (1964), pp. 40–63, **113** (1964), pp. 64–86; **221** (1976), pp. 1–33; **221** (1976), pp. 35–58; see also his review article in *Group representations in Mathematics and Physics*, Lecture Notes in Physics, No. 6, Springer-Verlag, Berlin, 1970.

CHAPTER VIII

KINEMATICS AND DYNAMICS

1. THE ABSTRACT SCHRÖDINGER EQUATION

In this chapter we shall use the main results of Chapters V through VII to study the dynamical and kinematical properties of quantum mechanical systems. This will lead to the Schrödinger equation, the well known formulas for the position, momentum, and spin observables, and to the standard expressions for the Hamiltonians which enter the quantum theory of the atom.

The first problem we study is that of determining the dynamical group of a system, say $\mathfrak{S}$. According to our general principles, we must exhibit the dynamical group as a one-parameter group of convex automorphisms of the convex set of all states of $\mathfrak{S}$. In this section we obtain the form of the most general dynamical group of a system whose logic is standard. Our analysis leads to the so-called abstract Schrödinger equation of motion. This equation, in its usual form, gives a set of directions for computing the possible values of the energy observable and describes the development of the state of the system in time. In keeping with the spirit of our development, we can say that the Schrödinger equation gives the *infinitesimal form* of the temporal development of an arbitrary quantum mechanical system.

Let us then consider a system $\mathfrak{S}$; the quantum mechanical character of $\mathfrak{S}$ is put in evidence by the assumption that the logic of $\mathfrak{S}$ is standard. We shall actually assume that the logic of $\mathfrak{S}$ is that associated with the complex number field. We write $\mathscr{H}$ for a separable infinite dimensional Hilbert space over $\mathbf{C}$ whose logic $\mathscr{L}$ is isomorphic to that of $\mathfrak{S}$. We identify $\mathscr{L}$ with the logic of $\mathfrak{S}$. We write $\mathscr{S}$ for the convex set of all states of $\mathscr{L}$.

According to our earlier discussion in Chapter III, the dynamical group of $\mathfrak{S}$ is then to be described by a Borel one-parameter group $t \to D(t)$ of convex automorphisms of $\mathscr{S}$, the Borel structure being such as to make the maps

$$t \to (D(t)p)(a)$$

Borel, for all $p \in \mathscr{S}$ and $a \in \mathscr{L}$. Now, by Gleason's theorem (cf. Chapter IV), $A \to p_A$ is a convex isomorphism of the set of von Neumann operators of trace 1 onto $\mathscr{S}$, p_A being the state $a \to \mathrm{tr}(P^a A)$ $(a \in \mathscr{L})$. Consequently, each $D(t)$ induces an automorphism of the convex set of von Neumann operators of trace 1, so that we can select, for each t, a symmetry U_t of $\mathscr{H}$ such that

$$D(t)p_A = p_{U_t A U_t^{-1}} \tag{1}$$

(cf. theorem 4.33). As $D(t) = D(t/2)^2$, it follows that each U_t is a unitary operator. Now, the map $t \to U_t$ need not have any measurability properties, as the choice of U_t, for each t, was made quite arbitrarily. But, if u and v are unit vectors of $\mathscr{H}$, and if we write A for the projection on the space spanned by u, and L for the space spanned by v,

$$p_{U_t A U_t^{-1}}(L) = |\langle U_t u, v\rangle|^2.$$

Consequently, the map $t \to |\langle U_t u, v\rangle|^2$ is Borel for all $u, v \in \mathscr{H}$. If we therefore write π for the canonical homomorphism of the unitary group of $\mathscr{H}$ onto the projective group of $\mathscr{H}$, this implies (corollary 7.2) that the map $t \to \pi(U_t)$ is a Borel homomorphism of $\mathbf{R}^1$ into the projective group. Consequently it is induced by some projective representation of $\mathbf{R}^1$ in $\mathscr{H}$ (cf. Chapter VII). Now, according to theorem 7.38, all multipliers of $\mathbf{R}^1$ are exact. Hence there exists a (unitary) representation V

$$V : t \to V_t \tag{2}$$

of $\mathbf{R}^1$ in $\mathscr{H}$, such that for each t, and each von Neumann operator A of trace 1, $D(t)$ is given by

$$D(t) : p_A \to p_{V_t A V_t^{-1}}. \tag{3}$$

The representation V is determined up to multiplication by a character of the real line, i.e., if $V'(t \to V_t')$ is another representation satisfying (3), then there is a real number c such that

$$V_t' = e^{itc} V_t \tag{4}$$

for all t.

To the one-parameter group V we can apply Stone's theorem (cf. Riesz-Nagy [1]). According to this theorem, there exists a unique self-adjoint operator H such that

$$V_t = \exp(-itH) \tag{5}$$

for all t; the domain $\mathscr{D}_H$ of H is given by

$$\mathscr{D}_H = \left\{ v : \lim_{t \to 0} \frac{1}{t} (V_t v - v) \text{ exists} \right\}, \tag{6}$$

and, for $v \in \mathscr{D}_H$,

$$Hv = i \lim_{t \to 0} \frac{1}{t} (V_t v - v). \tag{7}$$

We can thus say, in view of (5) and (3), that H *induces* the dynamical group $t \to D(t)$. Since the representation V is determined only up to multiplication by a character of the real line, D determines H up to an additive constant, i.e., H and $H + c \cdot 1$ (c real) induce the same dynamical group. Collecting these facts, we have the following theorem:

Theorem 8.1. *Given any self-adjoint operator H on the complex separable Hilbert space $\mathscr{H}$ of infinite dimension, there exists a unique dynamical group $D(t \to D(t))$ of $\mathscr{S}$ such that, for each t, $D(t)$ transforms the state corresponding to the von Neumann operator A into the state corresponding to the von Neumann operator* $\exp(-itH)A \exp(itH)$. *H and $H + c \cdot 1$ (c a real number) induce the same dynamical group. Conversely, any dynamical group $D(t \to D(t))$ of convex automorphisms of $\mathscr{S}$ is induced in the above fashion by a self-adjoint operator H, and D determines H up to an additive constant. Finally, for any t, $D(t)$ transforms the pure state determined by the unit vector u into the pure state determined by the unit vector* $\exp(-itH)u$.

Theorem 11.1 is the central result of elementary quantum theory. It tells us that the dynamical groups of physical systems correspond one-one to one-parameter unitary groups and even to self-adjoint operators, provided we ignore additive constants in the self-adjoint operators. Now, according to our basic principles, the observables of the system are in one-one correspondence with the self-adjoint operators of $\mathscr{H}$. Consequently, we reach the remarkable conclusion that to every dynamical group there can be associated, in a canonical fashion, a class of observables, any two of which differ by an additive constant, and that any such class will generate a dynamical group. This conclusion brings us to a very interesting formal analogy with classical mechanics. In its Hamiltonian form (Chapter I, Volume I), the dynamical group of a classical system is determined also by a function H on the phase space, i.e., a classical observable, and, since H enters the differential equations in terms of its first-order derivatives, this observable in turn is determined by the dynamical group only up to an additive constant.

Prompted by this analogy, we shall call any self-adjoint operator H, which induces the dynamical group of a system, a *Hamiltonian of the system.* Given a physical system, the observable which corresponds to a Hamiltonian of the system is undoubtedly a very crucial one, since it describes the dynamics of the system completely. In physics, this observable is usually identified with the *energy* of the system. We shall give, further on, some of the reasons that make this identification a very reasonable one.

Assuming this identification, we can now recover the classical prescription of Schrödinger for the determination of the energy levels of the system. If u is a unit vector of $\mathscr{H}$ which describes a pure state of the system under consideration, and if $P(E \to P_E)$ is the projection valued measure giving the spectral decomposition of the operator H, then the probability distribution of energy, when the system is in the state corresponding to u, is the measure

$$E \to \|P_E u\|^2; \tag{8}$$

if H has a purely discrete simple spectrum, say $\{\lambda_1, \lambda_2, \cdots\}$, and if $u_1, u_2, \cdots$ are orthonormal eigenvectors of H corresponding to the eigenvalues $\lambda_1, \lambda_2, \cdots$, respectively, then $\lambda_1, \lambda_2, \cdots$ are the possible values of the energy observable, say En, and, in the given state,

$$\Pr(En = \lambda_m) = |\langle u_m, u\rangle|^2. \tag{9}$$

Let us now consider the dynamical group induced by the unitary group $t \to \exp(-itH)$. It is well known, and easy to prove in the context of Hilbert spaces (cf. Gel'fand [1]), that there are vectors u in $\mathscr{H}$ such that $t \to \exp(-itH)u$ is an analytic function of t with values in $\mathscr{H}$; and that, moreover, such vectors form a linear manifold which is dense in $\mathscr{H}$, and on which H is essentially self-adjoint. If we take such a vector u and write $u_t = \exp(-itH)u$, then we obtain the differential equation

$$\frac{du_t}{dt} = -iHu_t, \qquad u_0 = u. \tag{10}$$

The group $t \to \exp(-itH)$ can be clearly recovered from (10). The equation (10) is thus the differential equation which describes the *motion* of the pure states of our system. We shall call it the *abstract Schrödinger equation* (cf. Dirac [1], pp. 108–111).

From the representation $t \to V_t$ we can now determine how the probability distribution of any observable changes with time. Let ξ be any observable and A the self-adjoint operator associated with it. Let $E \to P_E$ be the spectral measure of A. If the unit vector u_0 represents the state at time $t=0$, then the distribution of ξ at time t is the measure q_t where

$$q_t(E) = \langle \exp(itH)P_E \exp(-itH)u_0, u_0\rangle. \tag{11}$$

It is of course quite possible that the distribution of ξ at time $t=0$ remains unchanged for all time, no matter what the initial state is. From (11) we infer that the necessary and sufficient condition for this is that

$$V_t^{-1} P_E V_t = P_E$$

for all t and all Borel sets E, i.e.,

$$\exp(-itH)\exp(-isA) = \exp(-isA)\exp(-itH) \tag{12}$$

for all s and t. The equation (12) yields, on differentiation,

$$[H,A] = 0, \tag{13}$$

where (13) is interpreted in the strong sense, i.e., the spectral projections of H and A commute. In other words, the observable ξ must be simultaneously observable with energy. In analogy with classical mechanics we shall call ξ an *integral of the motion*. The reader may notice that (13) is also the condition for ξ to be an integral of the classical motion determined by a classical Hamiltonian H, provided we interpret the bracket in (13) as the Poisson bracket. Later on, we shall come across further instances of this analogy. *In particular, energy is always conserved during motion.*

We shall conclude this section with a few remarks on the "Heisenberg picture." Consider a system whose dynamical development is described by the one-parameter group $t \to V_t$. The pure states of the system then move in time along the parametrized curves $t \to V_t u$. We shall call these curves the *motions* of the system or, alternately, the *dynamical states* of the system. Clearly any such motion $t \to u_t$ is completely determined by the vector u_0. Hence the rays of the Hilbert space of the system can also be regarded as representing the dynamical states of the system. If u and v are two unit vectors which represent the dynamical states ξ and η, respectively, the number $|\langle u,v\rangle|^2$ can evidently be interpreted as the probability of finding the system in the dynamical state η when a measurement is made on it to find out whether it is in the dynamical state ξ or not. Thus $|\langle u,v\rangle|^2$, which depends only on the rays determined by u and v, has an invariant physical significance.

It is possible to reformulate the general prescriptions of quantum mechanics in terms of the concept of the dynamical state. We shall indicate how this is done by computing the probability distributions of the observables as time changes. Suppose x is an observable and $A_x = A$ its operator. Suppose that the system is in the dynamical state u_0. Then, the probability distribution of x at time t is the measure

$$q_t : E \to \langle P_E V_t u_0, V_t u_0\rangle,$$

where $E \to P_E$ is the spectral measure of A. If we write $A_t = V_t^{-1} A V_t$, then $P^t : E \to V_t^{-1} P_E V_t$ is the spectral measure of A_t and

$$q_t(E) = \langle P_E^t u_0, u_0\rangle,$$

which describes q_t entirely in terms of u_0. We see, therefore, that the probability distribution of x at time t can be computed if we replace A by A_t, but evaluate the relevant inner products with u_0 instead of u_t. In other words, *we may regard the change in time as inducing a change in the operators which represent x; the dynamical state remains unchanged.*

It must be realized, however, that this is only an alternative *language* to describe the same state of affairs. In physical literature, this way of looking at things is called the *Heisenberg picture* and the dynamical states are called the *Heisenberg states*. Our original way of looking at things is referred to as the *Schrödinger picture.*

If A is the operator which represents the observable x, then at time t, it will be represented by the operator $A_t = V_t^{-1} A V_t$ in the Heisenberg picture. The (formal) time derivative of A_t at $t=0$ then becomes $i[H,A]$, where we write $V_t = \exp(-itH)$:

$$\dot{A} = i[H,A]. \tag{14}$$

We thus obtain the celebrated principle of Heisenberg according to which the time derivative of an observable is i times the commutator of its operator (at time 0) with the energy. Appropriately interpreted, (14) retains its meaning in classical mechanics also.

2. COVARIANCE AND COMMUTATION RULES

We shall now consider a system such as a particle. A characteristic property of such systems is that one can associate with each of them, a certain C^∞ manifold M, which is usually called its configuration space. If we are interested in a *classical* interpretation of such a system, M would *actually* be its configuration space. Suppose that $\mathfrak{S}$ is such a system and that we are interested in a quantum mechanical description of $\mathfrak{S}$. Suppose we still like to treat the configuration of the system as a physical observable. The mathematical way to put this in evidence is to assign, for each Borel set $E \subseteq M$, a closed linear manifold $S(E)$ of the Hilbert space $\mathscr{H}$ of $\mathfrak{S}$; $S(E)$ will then represent the experimental statement that the configuration of $\mathfrak{S}$ belongs to E. It is, moreover, natural to assume that the assignment $E \rightarrow S(E)$ is a σ-homomorphism of the Borel structure of M onto a Boolean sub-σ-algebra of the logic $\mathscr{L}$ of $\mathscr{H}$, i.e., that the mapping $E \rightarrow P^{S(E)} = P_E$, which assigns to E the projection on $S(E)$, is a projection valued measure based on M.

Let us now suppose that we have in addition a Lie group G acting differentiably on M and that we want a description of our system which is covariant with respect to G (for example, we may treat a particle in a spherically symmetric potential field in space; M will then be R^3 and G will be the rotation group). We then associate, with each element g of G, an automorphism α_g of the logic $\mathscr{L}$ of our system such that $\alpha(g \rightarrow \alpha_g)$ is a representation of G in $\mathrm{Aut}(\mathscr{L})$; moreover, in view of the simple meaning of the elements $S(E)$ of $\mathscr{L}$, we shall require that α and S be related as follows:

$$\alpha_g[S(E)] = S(g[E]). \tag{15}$$

We take the point of view that (15) summarizes the physical assumptions we have made. If G is a connected Lie group, we can select a projective representation $U(g \to U_g)$ of G in $\mathscr{H}$ which induces α. Then (15) can be written as

$$U_g P_E U_g^{-1} = P_{g[E]}. \tag{16}$$

If U is a representation, then (16) expresses the fact that (U,P) is a system of imprimitivity. If U were not a representation but its multiplier is exact, then we may change U to a representation so that it still induces α. In such a case, we may therefore say that the existence of a configuration space plus the requirement of covariance leads to a system of imprimitivity, the representation and projection valued measure of which have direct physical significance. The exactness of the multipliers of the projective representations which induce α cannot of course be always taken for granted. But in many cases this is the case (cf. theorem 7.40).

The equations (16) can be written formally in their infinitesimal forms. Let U be a fixed representation satisfying (16). Let $\mathfrak{g}$ be the Lie algebra of G. For any $X \in \mathfrak{g}$ and each t, $x \to \exp(tX)\cdot x$ is a diffeomorphism of M; and as t varies, we obtain a one-parameter group. Let τ_X be the vector field of this one-parameter group. For any $f \in C^\infty(M)$ and $x \in M$, we have (by definition),

$$(\tau_X f)(x) = \frac{d}{dt}(f(\exp(tX)\cdot x))_{t=0}. \tag{17}$$

For any real C^∞ function f on M, let A_f be the operator

$$A_f = \int_M f\,dP. \tag{18}$$

A_f is possibly unbounded but self-adjoint. Then, (16) leads to

$$U_g A_f U_g^{-1} = A_{(f^g)}. \tag{19}$$

Now, $t \to U_{(\exp tX)}$ is a representation of R^1, and hence there exists a self-adjoint operator B_X such that $\exp(-itB_X) = U_{(\exp tX)}$ for all t. B_X is determined uniquely and we have, formally from (19),

$$\begin{aligned} [A_f, A_{f'}] &= 0, \\ [A_f, B_X] &= iA_{(\tau_X f)}, \\ [B_X, B_Y] &= iB_{[X,Y]}, \\ aB_X + bB_Y &= B_{aX+bY}. \end{aligned} \tag{20}$$

We shall call the equations (20) the *Heisenberg commutation rules* (cf. (10), Chapter I).

The observables associated with the operators A_f and B_X which enter the relations (20) may be given interesting physical interpretations. We consider first the operators A_f. We recall that for any set E, $S(E)$ represents the experimental statement that the configuration of $\mathfrak{S}$ belongs to E. Now, from the formula (18) for A_f, we see that the operator A_f has the spectral measure $F \rightarrow P_{f^{-1}(F)}$ (F a Borel set of the line). Hence, if we denote by ζ the observable corresponding to A_f, the statement that the value of ζ lies in F must be represented by the subspace $S(f^{-1}(F))$. It is thus identical with the statement that the configuration of $\mathfrak{S}$ belongs to the set $f^{-1}(F)$. In other words, if we observe, instead of the configuration x of $\mathfrak{S}$, the value of f at x, we obtain an observable whose associated self-adjoint operator is A_f. We shall therefore call these observables *configuration observables*. Note that the function f represents classically an observable with the same significance. Thus $f \rightarrow A_f$ is an algebra homomorphism which assigns to each classical configuration observable the "corresponding" quantum observable. For example, if $M = R^n$ and f is the kth coordinate function $(t_1, \cdots, t_n) \rightarrow t_k$, then A_f is the operator which represents the kth configuration coordinate.

On the other hand, the vector field τ_X gives rise to the classical momentum observable corresponding to the symmetries $t \rightarrow \exp(tX)$ acting on M, and the analogies we have so far developed with classical mechanics lead us to define B_X as the quantum mechanical *momentum observable* corresponding to the symmetries generated by X. In other words, the covariance expressed by (20) can be formally regarded as describing certain commutation rules between the configuration observables and the momentum observables of the system. We have thus arrived at the real mathematical meaning of the celebrated commutation rules of physics; at the same time our extensive knowledge of systems of imprimitivity enables us to understand why they are so far-reaching.

3. THE SCHRÖDINGER REPRESENTATION

Given M and G, the main problem of course is to determine the pairs (α, S) which satisfy the relations (15). This is, at least in the special case when α is induced by a representation of G, nothing but the problem of determining the systems of imprimitivity for G which are based on M. The analysis of Chapter VI shows that such a determination is completely known when G acts transitively on M. Fortunately, in many physical applications, G *is* transitive on M, so that these results can be applied. On the other hand, whether or not G acts transitively on M, there always exist the Koopman systems of imprimitivity associated with the invariant measure classes on M. In this section we shall explicitly describe these latter systems in the form in which they are most familiar in physics.

We start with the C^∞ manifold M and a Borel measure ν on it. We write $\mathscr{H} = \mathscr{L}^2(M,\nu)$. For any real C^∞ function f on M we define A_f to be the operator of multiplication by f:

$$A_f : \varphi \to f\varphi. \tag{21}$$

For any C^∞ vector field Z on M we define B_Z as the operator

$$B_Z : \varphi \to -iZ\varphi. \tag{22}$$

Both (21) and (22) are meaningful only when φ is suitably restricted, certainly when $\varphi \in C_c^\infty(M)$. Now, if we define A_f rigorously by (21) on the set of all φ such that φ and $f\varphi$ are both in $\mathscr{H}$, then it is well known that A_f is self-adjoint and is the closure of its restriction to $C_c^\infty(M)$. On the other hand, it is not always possible to make B_Z self-adjoint even if it is formally symmetric on $C_c^\infty(M)$; the well known example is of course the one where $M =]0,\infty[$, and $Z = d/dx$ (ν is Lebesgue measure). We shall now formulate the following useful lemma which describes a set of sufficient conditions for B_Z to be essentially self-adjoint when defined by (22) on $C_c^\infty(M)$.

Lemma 8.2 (*van Hove* [1]). *Let $t \to g_t$ be a one-parameter group of diffeomorphisms of M such that $t, x \to t \cdot x = g_t(x)$ is C^∞ from $\mathbf{R}^1 \times M$ into M. Let Z be the vector field induced by $t \to g_t$. Suppose that ν is invariant under all the g_t. Then, the operator B_Z is symmetric on the domain $C_c^\infty(M)$, and its closure is self-adjoint in $\mathscr{H}$. If we denote this closure by B_Z also, then for any t,* $\exp(-itB_Z)$ *is the operator $f \to f^{(g_t)}$ in $\mathscr{H}$, i.e.,*

$$(\exp(-itB_Z)f)(x) = f(g_t^{-1}x). \tag{23}$$

Proof. By definition, $(Z\varphi)(x) = ((d/dt)\varphi(t \cdot x))_{t=0}$ for all $x \in M$ and $\varphi \in C_c^\infty(M)$. We begin our proof with the observation that a complex valued Borel measure ν' is invariant with respect to the g_t if and only if

$$\int_M (Z\varphi)d\nu' = 0 \tag{24}$$

for all $\varphi \in C_c^\infty(M)$. In fact, let $u(s) = \int_M \varphi(s \cdot x)d\nu'(x)$. If ν' is invariant, $u(s)$ is independent of s and hence has vanishing derivative, proving (24). On the other hand, if u is as above, we have,

$$\begin{aligned}\frac{du}{ds} &= \int_M (Z\varphi)(s \cdot x)d\nu'(x),\\ &= \int_M (Z(\varphi^{-s}))(x)d\nu'(x),\end{aligned}$$

so that (24) implies that u is a constant, giving the invariance of ν'. This said, we come to the proof proper. For $\varphi, \psi \in C_c^\infty(M)$,

$$\int_M \varphi(t\cdot x)\psi(x)d\nu(x) = \int_M \varphi(x)\psi(t^{-1}\cdot x)d\nu(x),$$

since ν is invariant. Therefore, we have, on differentiation,

$$\int_M (Z\varphi)\psi d\nu = -\int_M \varphi(Z\psi)d\nu,$$

and it follows that the operator $B_Z : \varphi \to -iZ\varphi$ is symmetric on $C_c^\infty(M)$. We shall now prove that the closure of B_Z in $\mathscr{H}$ is self-adjoint. It is known from standard spectral theory that a necessary and sufficient condition for this is that for any *nonreal* complex number ζ, the operator $B_Z+\zeta\cdot 1$ map $C_c^\infty(M)$ onto a dense linear manifold of $\mathscr{H}$. Suppose that $f_0 \in \mathscr{H}$ and

$$-i\int_M (Z\varphi)f_0 d\nu + \zeta\int_M \varphi f_0 d\nu = 0$$

for all φ in $C_c^\infty(M)$. We must prove that $f_0=0$. Let $\rho=\zeta/i$. Then ρ is not purely imaginary and

$$\int_M (Z\varphi)f_0 d\nu = \rho\int_M \varphi f_0 d\nu \tag{25}$$

for all $\varphi \in C_c^\infty(M)$.

Consider now the manifold $M'=\mathbf{R}^1\times M$ and the one-parameter group $t\to g_t'$ of diffeomorphisms of M', where

$$g_t'(s,x) = (t+s, t\cdot x).$$

Write Z' for the corresponding vector field; for $u \in C_c^\infty(\mathbf{R}^1)$ and $\varphi \in C_c^\infty(M)$, let $u\otimes\varphi$ be the function $s, x \to u(s)\varphi(x)$. Then,

$$\begin{aligned}\iint_{M'} (Z'(u\otimes\varphi))(s,x)e^{\rho s}f_0(x)d\nu(x)ds &= \int_{-\infty}^{\infty}\frac{du}{ds}(s)e^{\rho s}ds\int_M \varphi(x)f_0(x)d\nu(x)\\ &\quad + \int_{-\infty}^{\infty} u(s)e^{\rho s}ds\int_M (Z\varphi)(x)f_0(x)d\nu(x)\\ &= 0,\end{aligned}$$

as is easily seen by integrating by parts the integral with respect to ds in the first term and using (25). A standard approximation argument (cf. Schwartz [1], pp. 108–109) now leads to

$$\iint_{M'} (Z'w)(s,x)e^{\rho s}f_0(x)ds d\nu(x) = 0$$

for all $w \in C_c^\infty(M')$. In view of our observation at the beginning of the proof (cf. (24)) we conclude that the complex Borel measure λ, where

$$\lambda(E) = \iint_E e^{\rho s} f_0(x) ds d\nu(x),$$

is invariant with respect to the group $t \to g_t'$. But $dsd\nu$ is also an invariant measure. Hence for each t, we must have $e^{\rho(s+t)} f_0(t\cdot x) = e^{\rho s} f_0(x)$ for almost all s, x, i.e.,

$$f_0(t\cdot x) = e^{-\rho t} f_0(x)$$

for ν-almost all x. Therefore, for each real t,

$$\int_M |f_0(x)|^2 d\nu(x) = \left(\int_M |f_0(x)|^2 d\nu(x)\right) |e^{-\rho t}|^2,$$

and since ρ is not purely imaginary, we conclude that $f_0 = 0$. As mentioned before, this proves that the closure of the operator $\varphi \to -iZ\varphi$ is self-adjoint.

Now the operators V_t, where

$$(V_t f)(x) = f(t^{-1}\cdot x) \qquad (f \in \mathscr{H}),$$

are unitary and $V_t = \exp(-itA)$ for a self-adjoint operator A. Further, $A\varphi = B_Z\varphi$ whenever $\varphi \in C_c^\infty(M)$. Hence A must coincide with the closure of B_Z. This finishes the proof of the lemma.

Consider now a connected Lie group G acting on M such that the map $g, x \to g\cdot x$ of $G \times M$ into M is C^∞. Let ν be a Borel measure invariant under G. We then take $\mathscr{H} = \mathscr{L}^2(M,\nu)$ and define

$$\begin{aligned} P_E\varphi &= \chi_E \cdot \varphi, \\ (U_g\varphi)(x) &= \varphi(g^{-1}x). \end{aligned} \tag{26}$$

(U,P) is a Koopman system of imprimitivity associated with ν and G. If we take $\mathscr{H}$ as the Hilbert space of a system with M as its configuration space and G as its group of symmetries, then we see that the configuration observables are represented by the multiplication operators A_f ($f \in C^\infty(M)$). For any element $X \in \mathfrak{g}$, the Lie algebra of G, the momentum observable B_X corresponding to the symmetries $x \to \exp(tX)\cdot x$ of M is the differential operator $-i\tau_X$, where τ_X is the vector field on M defined by

$$(\tau_X\varphi)(x) = \left(\frac{d}{dt}\varphi(\exp(tX)\cdot x)\right)_{t=0};$$

the differential operator $B_X = -i\tau_X$ is the self-adjoint operator which is the closure of the operator $\varphi \to -i\tau_X\varphi$ on $C_c^{\infty}(M)$ and $U_{(\exp tX)} = \exp(-itB_X)$ for all t. The B_X and A_f satisfy the commutation rules:

$$\begin{aligned} [A_f, A_{f'}] &= 0, \\ [B_X, A_f] &= iA_{(\tau_X f)}, \\ [B_X, B_Y] &= iB_{[X,Y]}, \\ aB_X + bB_Y &= B_{(aX+bY)}, \end{aligned} \tag{27}$$

where $f, f' \in C^{\infty}(M)$, a, b are constants, and X, $Y \in \mathfrak{g}$; the equations (27) are valid on the dense linear manifold $C_c^{\infty}(M)$.

This particular representation of the equations (20) is known as the *Schrödinger representation*. From our point of view the relations (16) are the fundamental ones; the commutation rules (20) are to be regarded merely as their formal infinitesimal versions. The reason for the frequent appearance of the commutation rules in physical literature is simply due to the tradition in mechanics and physics according to which the laws governing physical systems were always expressed in the infinitesimal form whenever possible. Rigorous formulations of (20) can be given; but they get involved with the domain considerations inevitable in any considerations of unbounded operators, and the equations lose the global geometric flavor possessed by the "integrated" relations (16). The mathematical importance of the integrated versions of (20) seems to have been first recognized by Weyl ([1], pp. 272–276).

When G acts transitively and freely on M, we may use theorem 9.17 to conclude that every irreducible system of imprimitivity for G based on M is equivalent to the Koopman system. We may state this loosely by saying that every irreducible representation of the commutation rules (20) is equivalent to the Schrödinger representation. When M was R^n and G was the group of its translations, this was first proved by Stone [3] and von Neumann [3]. Our method of proof, based on theorem 6.17, is that of Mackey [9]. This result means that the apparently special choices made in the Schrödinger representation, for representing the states and the position and momentum observables, are in fact the most general ones possible subject to the commutation rules (20), provided we assume irreducibility. We may thus regard this theorem as proving the isomorphism of Schrödinger's wave mechanics (based on the Schrödinger representation), and Heisenberg's matrix mechanics (based on the commutation rules (20) only and not on any specific choice of the A_f and B_X), and thereby unifying the two great developments.

The irreducibility of the (U,P) can be given an interesting interpretation. The system (U,P) is irreducible if and only if the elements in the commuting ring of (U,P) are the multiples of the identity or if and only if the smallest von Neumann algebra containing the spectral projections of the B_X and A_f is the ring of all operators in $\mathscr{H}$. Let us define a self-adjoint operator A to be a *function of the* B_X *and* A_f if A belongs to the von Neumann algebra generated by the B_X and the A_f. Then every self-adjoint operator is a function of the B_X and the A_f. On the other hand, if we consider the classical system with configuration space M, it is true that all the observables are functions of the configuration observables and the linear momenta. We thus see that the assumption of reproducibility corresponds (roughly) to the assumption that there are essentially no new observables other than the configuration ones and the linear momenta.

4. AFFINE CONFIGURATION SPACES

An *affine space* is a pair (M, V), where (i) V is a real finite dimensional vector space, and (ii) V acts freely and transitively on M. By the usual abuse of language we shall refer to M itself as an affine space. V is said to be the vector space *associated with* M. For $v \in V$ we write t_v for the transformation of M induced by v and refer to it as the *translation by* v. We write

$$t_v \cdot x = x + v.$$

Given x and y in M, we shall write $y-x$ for the unique element v of V such that $t_v \cdot x = y$. Since t_{-v} and t_v are inverse to each other, $y-x = -(x-y)$, while the relation $t_{u+u'} = t_u t_{u'}$ leads to the equation $(x-y)+(y-z) = x-z$.

A one-one map L of M onto itself is called an *affine automorphism* of M if there exists a linear automorphism L^0 of the vector space V such that for all x and y in M,

$$Lx - Ly = L^0(x-y). \tag{28}$$

Since for fixed x, $y \to (x-y)$ maps M onto V, it is clear that L^0 is uniquely determined by L. If L^0 is the identity, we shall call L a *translation*. If $v \in V$, it is trivial to check that t_v is a translation. Conversely, let L be a translation and let x and y be two points such that $y = Lx$. If we write $y = x+v$, then it follows from the equation $Lz_1 - Lz_2 = z_1 - z_2$ that L is the translation $x \to x+v$. If we choose a point $x_0 \in M$ and map M bijectively onto V by the map $v \to x_0 + v$, then the affine automorphisms of M are easily seen to be precisely those which go over to the inhomogeneous transformations

$$a \to L^0 a + b \tag{29}$$

of V. Thus, in the notion of the affine automorphisms of M, we have an invariant description of the inhomogeneous linear transformations which are associated with vector spaces.

If we choose a basis $v_1, \cdots, v_n$ for V, and a point $x_0 \in M$, then the map

$$(t_1, \cdots, t_n) \rightarrow x_0 + \left(\sum_{j=1}^{n} t_j v_j \right) \tag{30}$$

is one-one from $\mathbf{R}^n$ onto M, and its inverse maps M onto $\mathbf{R}^n$ bijectively. We shall call the inverse of (30) an *affine coordinate system on M with origin in x_0*. Suppose $x_0, x_1 \in M$, $\{v_1, \cdots, v_n\}$, $\{w_1, \cdots, w_n\}$ are two bases for V, and γ, γ' the affine coordinate systems of M associated with x_0, $\{v_1, \cdots, v_n\}$ and x_1, $\{w_1, \cdots, w_n\}$. Then there exists $(a_1, \cdots, a_n) \in \mathbf{R}^n$ and an invertible matrix $(d_{ij})_{i,j=1}^n$ such that, for any $x \in M$, the images $(t_1, \cdots, t_n) = \gamma(x)$ and $(t_1', \cdots, t_n') = \gamma'(x)$ are related by

$$t_i' = \sum_{j=1}^{n} d_{ij} t_j + a_i. \tag{31}$$

We thus obtain the usual formulas which connect the two coordinate systems. From the formula (31) it follows that there is a unique C^∞ structure on M such that each affine coordinate system is a diffeomorphism of M onto $\mathbf{R}^n$. With this C^∞ structure, the map $v, x \rightarrow x + v$ is C^∞ from $V \times M$ into M.

It is obvious that M admits an invariant Borel measure, say dm. dm goes over to a multiple of the Lebesgue measure in $\mathbf{R}^n$ under any affine coordinate system. From our analysis in Section 2 we conclude that the only irreducibly covariant systems with M as their configuration space are described by the Schrödinger representation. The Hilbert space of such systems can be taken to be $\mathscr{L}^2(dm)$. The configuration observables are then represented by the multiplication operators. If $v \in V$, the momentum observable corresponding to the symmetries $x \rightarrow x + tv$ is represented by the operator B_v:

$$B_v = -i\partial(v), \tag{32}$$

where $\partial(v)$ is the usual linear differential operator associated with v. We shall call this the *linear momentum observable in the direction of the vector v*.

If we choose now a point $x_0 \in M$, a basis $\{v_1, \cdots, v_n\}$ of V, and map M onto $\mathbf{R}^n$ with the help of the corresponding affine coordinate system γ, we can obtain concrete forms for these results. For simplicity of description we identify M with $\mathbf{R}^n$ and $\mathscr{H}$ with $\mathscr{L}^2(\mathbf{R}^n)$. Then the kth coordinate function, say q_k, is represented by the operator of multiplication by t_k:

$$q_k \sim \varphi(t_1, \cdots, t_n) \rightarrow t_k \varphi(t_1, \cdots, t_n), \tag{33}$$

and the linear momentum in the direction of the jth coordinate axis, say p_j, is represented by the operator $-i(\partial/\partial t_j)$:

$$p_j \sim \varphi(t_1, \cdots, t_n) \rightarrow -i \frac{\partial}{\partial t_j} \varphi(t_1, \cdots, t_n). \tag{34}$$

The second of the commutation rules (20) then becomes

$$[q_k,p_j] = (-1)^{1/2}\delta_{kj}1, \tag{35}$$

which is the form in which the rules were stated first by Heisenberg.

If φ is a unit vector in $\mathscr{H}$, i.e.,

$$\int_M |\varphi(x)|^2 dm(x) = 1, \tag{36}$$

then the probability distribution q_φ of the configuration of the system in the (pure) state represented by φ is given by

$$q_\varphi(E) = \int_E |\varphi(x)|^2 dm(x). \tag{37}$$

Since we are identifying M with $\mathbf{R}^n$ using a coordinate system, the formula (37) becomes

$$q_\varphi(E) = \int_E |\varphi(t_1,\cdots,t_n)|^2 dt_1\cdots dt_n. \tag{38}$$

The formulas (34) and (35) were first postulated by Schrödinger when M was three dimensional and the system to be analyzed was that of a single particle. Formula (34) gave the forms of the *position* operators, (35) the forms of the components of the *linear momenta.* Formula (38) gave the probability that in the state represented by φ, called the *wave function of the particle*, the particle is to be found in the region E and was first recognized by Max Born. From our point of view, we see that the whole development falls out of the assumption of affine covariance and the irreducibility of the system (U,P).

The special forms of the configuration and momentum observables can also be used to illustrate the phenomenon of *complementarity* so peculiar to quantum theory. Suppose $\mathscr{H}$, U, P are as above. Let us choose an affine coordinate system for M and identify M with $\mathbf{R}^n$ and $\mathscr{H}$ with $\mathscr{L}^2(\mathbf{R}^n)$ via this coordinate system. The expressions for the configuration observables show that they commute among themselves. On the other hand, the translation operators U_v also commute among themselves, so that the momentum operators B_v commute among themselves. Let us now take a function $\varphi \in \mathscr{L}^2(\mathbf{R}^n)$ of unit norm and compute the probability distributions of the vector observables $\mathbf{q}=(q_1,\cdots,q_n)$ and $\mathbf{p}=(p_1,\cdots,p_n)$. For $\mathbf{q}$, this is straightforward; it is

$$E \to \int_E |\varphi(t_1,\cdots,t_n)|^2 dt_1\cdots dt_n. \tag{39}$$

To compute the distribution of $\mathbf{p}$ we must use Fourier transforms. Let

$$\mathscr{F} : \psi \to \hat{\psi} \tag{40}$$

be the Fourier-Plancherel transform of $\mathscr{L}^2(\mathbf{R}^n)$ onto itself which is the unique unitary automorphism of $\mathscr{L}^2(\mathbf{R}^n)$ with the property that for all $\psi \in C_c(\mathbf{R}^n)$,

$$\hat{\psi}(t_1, \cdots, t_n) = \frac{1}{(2\pi)^{n/2}} \int_{\mathbf{R}^n} \exp -i[t_1x_1 + \cdots + t_nx_n] \times \psi(x_1, \cdots, x_n)dx_1 \cdots dx_n \tag{41}$$

for all $(t_1, \cdots, t_n) \in \mathbf{R}^n$ (Bochner-Chandrasekharan [1]). It is then easy to show that for $v = (v_1, \cdots, v_n)$, $\mathscr{F} U_v \mathscr{F}^{-1}$ is the operator of multiplication by $\exp[-i(v_1t_1 + \cdots + v_nt_n)]$. From this it follows easily that the spectral measure of B_v is given by $F \to \mathscr{F}^{-1} P_{F_v} \mathscr{F}$, where

$$F_v = \{(t_1, \cdots, t_n) : v_1t_1 + \cdots + v_nt_n \in F\} \tag{42}$$

(F a Borel set on the line). From (42) we see that the probability distribution of $\mathbf{p}$ in the state given by φ is

$$E \to \int_E |\hat{\varphi}(t_1, \cdots, t_n)|^2 dt_1 \cdots dt_n. \tag{43}$$

Formulas (39) and (43) show the remarkable statistical connection between the configuration observables and momentum observables. If φ is such that φ vanishes outside of a small region S, the configuration of the system is localized in S statistically; formula (43) then tells us that the distribution of the momenta is very diffuse. Conversely, states φ for which $\hat{\varphi}$ has a small region for its support localize the momenta; but the configuration observables will have diffuse probability distributions.

No single state exists in which both the position and the momenta can be localized sharply and the same system exhibits both types of properties in appropriate states.

5. EUCLIDEAN SYSTEMS: SPIN

We shall now discuss the situation when G is transitive on M but its action is no longer free. The different systems with M as their configuration space and G as their symmetry group will now be labeled by representations of the stability groups. In the special case when M is a three-dimensional Euclidean affine space and G is the covering group of its group of motions, this leads to the kinematical classification of such systems in terms of their spin.

A *Euclidean affine space* is a triple $(M, V, \langle .,. \rangle)$, where M is an affine space with associated vector space V, and $\langle .,. \rangle$ is a positive definite inner product on $V \times V$. By the usual abuse of language we shall refer to M itself as a Euclidean space. An affine automorphism L of M is called a *motion* of M if the associated linear transformation L^0 of V is orthogonal with respect to $\langle .,. \rangle$ and has determinant $+1$. The set of all motions of M is a group, denoted by G. For each $v \in V$, the translation $t_v : x \to x+v$ belongs to G and $v \to t_v$ is an isomorphism of V onto a normal subgroup T_0 of G. For any $x \in M$ we write G_x for the stability subgroup of G at x. Given $x_0 \in M$ and a basis $\{v_1, \cdots, v_n\}$ ($n = \dim V$) of V, the affine coordinate system relative to these is called a *Euclidean coordinate system*, if $\{v_1, \cdots, v_n\}$ is an orthonormal basis for V. It is clear that if we map M onto $\mathbf{R}^n$ using a Euclidean coordinate system, G goes over to the inhomogeneous $SO(n)$-group in $\mathbf{R}^n$ and G_{x_0} goes over to the rotation group $SO(n)$, where x_0 is the origin of the coordinate system. From this observation we can obtain at once the structure of the group G. For any $x \in M$,

$$G = T_0 G_x, \qquad T_0 \cap G_x = \{\text{identity}\}. \tag{44}$$

Moreover, there is a canonical way to convert G into a Lie group. T_0 and G_x are closed, and G_x is compact. If $n \geq 2$, G is connected. If $n \geq 3$, G_x is semisimple and its universal covering group is also compact. We shall choose a universal covering group G^* over G and denote by δ the covering homomorphism of G^* onto G. δ is an isomorphism of T_0^*, the component of the identity of $\delta^{-1}(T_0)$, onto T_0; on the other hand, for any $x \in M$, $\delta^{-1}(G_x) = G_x^*$ is connected, simply connected, and covers G_x (through δ). G^* is a semidirect product of T_0^* and G_x^*.

Suppose we consider a quantum mechanical system $\mathfrak{S}$ for which a Euclidean affine space M is the configuration space. Assume the dimension of M to be ≥ 3. For a description of $\mathfrak{S}$ which is covariant with respect to the group G of motions of M, we must introduce a pair (U,P), where $P(E \to P_E)$ is a projection valued measure based on M and U a projective representation of G such that for all g and E,

$$U_g P_E U_g{}^{-1} = P_{g[E]}. \tag{45}$$

Both U and P act in $\mathscr{H}$, the Hilbert space of $\mathfrak{S}$. Now the group G satisfies the conditions of theorem 7.40 and hence we know that the universal covering group G^* of G has only exact multipliers. We may thus assume that the projective representation U arises from a unique unitary representation V of G^* in $\mathscr{H}$. V and P will then satisfy the relation

$$V_{g^*} P_E V_{g^*}{}^{-1} = P_{\delta(g^*)\cdot E} \qquad (g^* \in G^*). \tag{46}$$

The action $g^*, x \to \delta(g^*) \cdot x$ converts M into a transitive G^*-space and (46) tells us that (V,P) is a system of imprimitivity for G^*. (V,P) is irreducible if and only if the pair (U,P) is so. Conversely, suppose (V,P) is an irreducible system of imprimitivity for G^* based on M, acting in $\mathscr{H}$. Then the equation (46) shows that for any element g^* in the kernel of δ, V_{g^*} commutes with all P_E; but V_{g^*} also commutes with all the V_{x^*} ($x^* \in G^*$), since g^* lies in the center of G^*. Hence the irreducibility of (V,P) will imply that V maps the elements of the kernel of δ into scalars. From this we at once infer the existence of a projective representation U of G such that $x^* \to U_{\delta(x^*)}$ and $x^* \to V_{x^*}$ will define the same homomorphism of G^* in the projective group of $\mathscr{H}$.

Our problem is thus reduced to the determination of the irreducible systems of imprimitivity for G^*, based on M. Unlike the situation involving only the translations, there exist now irreducible systems (V,P) which are not unitarily equivalent. From the general theory developed in Chapter VI we know that the unitary equivalence classes of our pairs (V,P) are in canonical one-one correspondence with the equivalence classes of irreducible representations of the stability subgroup $G_{x_0}{}^*$ at some point $x_0 \in M$. However, the relation of unitary equivalence between the systems (V,P) is in general *finer* than the relation of physical equivalence; we shall call two systems of imprimitivity for G^* based on M, say (V,P) and (V',P'), acting in $\mathscr{H}$ and $\mathscr{H}'$, respectively, *physically equivalent* if there exists a unitary *or antiunitary* isomorphism W of $\mathscr{H}$ onto $\mathscr{H}'$ such that (i) $P = W \circ P' \circ W^{-1}$, and (ii) the representations V' and $W \circ V \circ W^{-1}$ of G^* in $\mathscr{H}'$ give rise to the same homomorphism into the projective group of $\mathscr{H}'$. Notice that as G^* has no one-dimensional representations (cf. theorem 7.40), (ii) is equivalent to requiring that $V' = W \circ V \circ W^{-1}$. If W is unitary we are back to unitary equivalence. But if W is antiunitary, the relation of physical equivalence will in general not imply unitary equivalence. We have, in fact, the following elementary result which clarifies the connection between unitary and physical equivalence.

Lemma 8.3. *Let X be a transitive Borel H-space for a lcsc group H and let $x_0 \in X$ be an arbitrary point. Let H_0 be the stability subgroup at x_0. Suppose that m_j $(j = 1, 2)$ are two unitary representations of H_0. Let (V^j,P^j) be the associated systems of imprimitivity acting in $\mathscr{H}^j$, say, $j = 1, 2$. Then, in order that there exist an antiunitary operator W of $\mathscr{H}^1$ onto $\mathscr{H}^2$ such that $V^2 = W \circ V^1 \circ W^{-1}$ and $P^2 = W \circ P^1 \circ W^{-1}$, it is necessary and sufficient that m_1 and m_2 be contragredient to each other.*

Proof. We choose a quasi-invariant measure α on X. Let $\mathscr{K}^j$ be the Hilbert space in which m_j acts. Then, according to the general theory of Chapter VI, there is a strict (H,X)-cocycle φ^j, whose values are in the unitary group of $\mathscr{K}^j$ with $(V^j,P^j) \simeq (\tilde{V}^j,\tilde{P}^j)$, where $(\tilde{V}^j,\tilde{P}^j)$ is associated with φ^j and α and acts in $\mathscr{L}^2(X,\mathscr{K}^j,\alpha)$. We choose a fixed (but otherwise arbitrary) antiunitary automorphism S of $\mathscr{K}^1$ and write $S^\sim$ for the mapping of $\mathscr{L}^2(X,\mathscr{K}^1,\alpha)$ onto itself defined by

$$(S^\sim f)(x) = Sf(x) \qquad (x \in X). \tag{47}$$

$S^{\sim}$ is evidently an antiunitary automorphism of $\mathscr{L}^2(X,\mathscr{K}^1,\alpha)$ and $S^{\sim}\tilde{P}^1 S^{\sim\,-1} = \tilde{P}^1$, as may be easily calculated from the formula $\tilde{P}_E{}^1 f = \chi_E f$. We write

$$\tilde{U}^1 = S^{\sim}\tilde{V}^1 S^{\sim\,-1}. \tag{48}$$

$(\tilde{U}^1,\tilde{P}^1)$ is also a system of imprimitivity for H based on X. Moreover, using the formula (39) of Chapter VI it can be easily calculated that $\tilde{U}^1$ is the system associated with the strict cocycle $\tilde{\varphi}^1$, where

$$\tilde{\varphi}^1(g,x) = S\varphi^1(g,x)S^{-1}. \tag{49}$$

Thus there exists an antiunitary isomorphism $\tilde{W}$ of $\mathscr{L}^2(X,\mathscr{K}^1,\alpha)$ onto $\mathscr{L}^2(X,\mathscr{K}^2,\alpha)$ such that $\tilde{W}\tilde{V}^1\tilde{W}^{-1} = \tilde{V}^2$ and $\tilde{W}\tilde{P}^1\tilde{W}^{-1} = \tilde{P}^2$ if and only if there exists a *unitary* isomorphism W' which sets up an equivalence between $(\tilde{U}^1,\tilde{P}^1)$ and $(\tilde{V}^2,\tilde{P}^2)$; in fact, W' and $\tilde{W}$ are related by the equation $W' = \tilde{W}S^{\sim\,-1}$. Hence a necessary and sufficient condition that there should exist an antiunitary operator W such that $W \circ V^1 \circ W^{-1} = V^2$ and $W \circ P^1 \circ W^{-1} = P^2$ is that the cocycles $\tilde{\varphi}^1$ and φ^2 define equivalent representations at x_0, i.e., the representations m_2 and $S \circ m_1 \circ S^{-1}$ are equivalent. Now there exists a canonical antiunitary isomorphism S_1 of $\mathscr{K}^1$ with its dual which intertwines m_1 and its contragredient $m_1{}^0$, i.e., $m_1 = S_1{}^{-1}m_1{}^0 S_1$. Thus our necessary and sufficient condition reduces to the equivalence of m_2 and $(SS_1{}^{-1}) \circ m_1{}^0 \circ (SS_1{}^{-1})^{-1}$, i.e., of m_2 and $m_1{}^0$, since $SS_1{}^{-1}$ is unitary. This proves the lemma.

Applying this lemma to the case with which we are concerned (since $G_{x_0}{}^*$ is compact, its irreducible representations are all finite dimensional), we conclude that the physical equivalence classes of irreducible systems (V,P) are in one-one correspondence with physical equivalence classes of representations of $G_{x_0}{}^*$, where two representations of $G_{x_0}{}^*$ are defined to be *physically equivalent* if they are either unitarily equivalent, or are contragredients of each other (it is easy to verify that this relation is an equivalence relation). In this sense, therefore, we have arrived at the complete solution to our problem.

All that remains is to obtain the explicit expressions for the systems of imprimitivity. We use dm for the measure on M which is invariant under the translations. It is obvious that dm is invariant under the motions of $G_{x_0}{}^*$ so that it is invariant under the action of the whole of G^*. To complete the construction, we must determine, for each irreducible representation $s(h \rightarrow s(h))$ of $G_{x_0}{}^*$, a strict (G^*,M)-cocycle φ^s with values in the unitary group of the Hilbert space $\mathscr{K}$ in which s acts, such that φ^s defines the representation s at x_0. We shall construct φ^s in the following way.

Since $T_0{}^*$ is normal in G^* and $G^* = T_0{}^* G_{x_0}{}^*$, we have, for any $g^* \in G^*$, the unique decomposition

$$g^* = t(g^*)r(g^*) \qquad (t(g^*) \in T_0{}^*,\ r(g^*) \in G_{x_0}{}^*). \tag{50}$$

The equation $g_1{}^* g_2{}^* = t(g_1{}^*)[r(g_1{}^*)t(g_2{}^*)r(g_1{}^*)^{-1}]r(g_1{}^*)r(g_2{}^*)$ shows at once that

$$r(g_1{}^* g_2{}^*) = r(g_1{}^*)r(g_2{}^*) \qquad (g_1{}^*,\ g_2{}^* \in G_{x_0}{}^*). \tag{51}$$

Moreover, for g^* in $G_{x_0}{}^*$,

$$r(g^*) = g^*. \tag{52}$$

(51) and (52) show that, if we define φ^s by

$$\varphi^s(g^*,x) = s(r(g^*)),$$

then φ^s is a strict cocycle of the type we want.

We are finally in a position to describe the irreducible pairs (V,P) explicitly. Let $s(h \to s(h))$ be an irreducible representation of $G_{x_0}{}^*$ in a (finite dimensional) Hilbert space $\mathscr{K}$ and let

$$\mathscr{H}^s = \mathscr{L}^2(M,\mathscr{K},dm). \tag{53}$$

For $E \subseteq M$ and $g^* \in G^*$ we define

$$\begin{aligned} P^s_E\psi &= \chi_E \cdot \psi, \\ (V^s_{g^*}\psi)(x) &= s(r(g^*))\psi(\delta(g^*)^{-1}\cdot x). \end{aligned} \tag{54}$$

Then (V^s,P^s) is the system of imprimitivity corresponding to the irreducible representation s of $G_{x_0}{}^*$.

The most important special case of the preceding discussion arises when M is three dimensional. We shall now proceed to consider this case in somewhat greater detail. We select a Euclidean coordinate system γ with origin x_0. Using γ, we shall identify M and V with $\mathbf{R}^3$; the bilinear form on $V \times V$ becomes the usual one. G will then be identified with the group of all rigid motions of $\mathbf{R}^3$. G_{x_0} goes over to the rotation group $SO(3)$. We write G_0 for G_{x_0} and $G_0{}^*$ for $G_{x_0}{}^*$.

We shall first construct an explicit covering of G_0 by $G_0{}^*$ following the well known construction of Weyl ([1], pp. 143–146). We introduce the Hermitian matrices (Pauli spin matrices)

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \tag{55}$$

and the correspondence $\mathbf{x} \to A(\mathbf{x})$, where $\mathbf{x} \in \mathbf{R}^3$ and

$$A(\mathbf{x}) = x_1\sigma_1 + x_2\sigma_2 + x_3\sigma_3. \tag{56}$$

Let

$$G_0{}^* = SU(2,\mathbf{C}). \tag{57}$$

Then for $g^* \in G_0{}^*$, $A(\mathbf{x}) \rightarrow (g^*)A(\mathbf{x})(g^*)^{-1}$ is a linear transformation of the vector space of all 2×2 Hermitian matrices of trace 0. Hence, via the correspondence (56), we obtain a linear transformation $\delta(g^*)$ of $\mathbf{R}^3$:

$$A(\delta(g^*)\mathbf{x}) = g^*A(\mathbf{x})g^{*-1}. \tag{58}$$

Since $\det A(\mathbf{x}) = -(x_1{}^2+x_2{}^2+x_3{}^2)$ is preserved under unitary similarities, $\delta(g^*)$ is an orthogonal transformation of $\mathbf{R}^3$. It is well known that $\delta(g^* \rightarrow \delta(g^*))$ is a covering homomorphism of $G_0{}^*$ onto G_0. Since the group G consists of elements $(g,\mathbf{v})$ $(\mathbf{v}\in \mathbf{R}^3, g \in G_0)$ with the composition

$$(g,\mathbf{v})(g',\mathbf{v}') = (gg',\mathbf{v}+g\cdot\mathbf{v}'), \tag{59}$$

we may define G^* to be $G_0{}^* \times \mathbf{R}^3$ with the composition

$$(g^*,\mathbf{v})(g^{*\prime},\mathbf{v}') = (g^*g^{*\prime},\mathbf{v}+\delta(g^*)\mathbf{v}'). \tag{60}$$

The map

$$\delta : (g^*,\mathbf{v}) \rightarrow (\delta(g^*),\mathbf{v}) \tag{61}$$

is a covering homomorphism from G^* onto G.

It is useful to describe the one-parameter subgroups of $G_0{}^*$ which map onto the familiar rotations around the axes. An easy computation leads to the following. If $t \rightarrow g_j{}^*(t)$ are the three one-parameter subgroups of $G_0{}^*$ defined by

$$g_3{}^*(t) = \begin{pmatrix} e^{it/2} & 0 \\ 0 & e^{-it/2} \end{pmatrix}, \quad g_1{}^*(t) = \begin{pmatrix} \cos t/2 & i \sin t/2 \\ i \sin t/2 & \cos t/2 \end{pmatrix},$$
$$g_2{}^*(t) = \begin{pmatrix} \cos t/2 & \sin t/2 \\ -\sin t/2 & \cos t/2 \end{pmatrix}, \tag{62}$$

then $t \rightarrow \delta(g_r{}^*(t)) = g_r(t)$ is given by

$$g_3(t) = \begin{pmatrix} \cos t & \sin t & 0 \\ -\sin t & \cos t & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad g_1(t) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos t & \sin t \\ 0 & -\sin t & \cos t \end{pmatrix},$$
$$g_2(t) = \begin{pmatrix} \cos t & 0 & -\sin t \\ 0 & 1 & 0 \\ \sin t & 0 & \cos t \end{pmatrix}. \tag{63}$$

The theory of the irreducible representations of $G_0{}^*$ is well known (cf. Weyl [1], pp. 143–180). The equivalence classes of these representations are labeled by a parameter j which takes the values $0, \frac{1}{2}, 1, \frac{3}{2}, \cdots$. For a given j, the associated representation, say D^j, is of dimension $2j+1$, and there exists an orthonormal basis in the representation space of D^j, say $\{v_{j\nu} : \nu = j, j-1, \cdots, -j\}$, such that

$$D^j(g_3{}^*(t))v_{j\nu} = e^{i\nu t}v_{j\nu}. \tag{64}$$

The basic representation $g^* \to g^*$ corresponds to $j = \frac{1}{2}$. The representations D^j for integral j map the kernel of δ $(=\{\pm 1\})$ into the identity, and hence induce representations of G_0. The other D^j induce only projective representations of G_0.

An explicit realization of the D^j can be obtained as follows. We regard $G_0{}^*$ as operating in the obvious manner on $\mathbf{C}^2 = W$ and consider the *tensor product* $W^{\otimes N} = W \otimes \cdots \otimes W$ (N factors, $N = 2j$). Then $W^{\otimes N}$ is a Hilbert space of dimension 2^N. If we define, for $g^* \in G_0{}^*$, $\pi(g^*)$ by

$$\pi(g^*) = g^* \otimes \cdots \otimes g^*, \tag{65}$$

then $\pi(g^* \to \pi(g^*))$ is a representation of G^* in $W^{\otimes N}$. π leaves the space of *symmetric* tensors invariant. Let W^j be this subspace. Then W^j is irreducible under π, and the representation defined by W^j is equivalent to D^j. If $e_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$, the reader can trivially check that

$$\pi(g_3{}^*(t))(e_1 \otimes \cdots \otimes e_1) = e^{ijt}(e_1 \otimes \cdots \otimes e_1). \tag{66}$$

Finally, as the representations D^j have different dimensions for distinct values of j, and as mutually contragredient representations have the same dimension, we see that each D^j is contragredient to itself. In view of our earlier discussion, the physical equivalence classes of the pairs (α, S) satisfying (15) depend on the parameter j $(j = 0, \frac{1}{2}, 1, \cdots)$, and for any j, (α, S) is the pair obtained from the corresponding system of imprimitivity induced by the representation D^j of $G_0{}^*$.

The Hilbert space of wave functions of a typical system which is irreducibly covariant under the Euclidean group of $\mathbf{R}^3$ can thus be regarded as the space $\mathscr{H}^j$ of square integrable Borel functions on $\mathbf{R}^3$ whose values are symmetric tensors of rank $(N = 2j)$ over the two-dimensional complex space W. Let W^j be the Hilbert space of symmetric tensors of rank $2j$ over W with norm $|\cdot|$ and inner product $\langle .,. \rangle$ induced naturally from W. Let D^j denote the representation of $G_0{}^*$ in W^j. Then, for the elements ψ of $\mathscr{H}^j$, we have

$$\int_{\mathbf{R}^3} |\psi(\mathbf{x})|^2 d\mathbf{x} < \infty, \tag{67}$$

and for any element $(g^*, \mathbf{v})$ of G^*, we have

$$(V_{(g^*,\mathbf{v})}\psi)(\mathbf{x}) = D^j(g^*)\psi(\delta(g^*)^{-1}\mathbf{x} - \delta(g^*)^{-1}\mathbf{v}), \tag{68}$$

while P_E is, as usual, the operator $\psi \rightarrow \chi_E \psi$. The operators of V induce the lattice automorphisms corresponding to the elements of G.

The formulas (67) and (68) are the usual ones in physical literature which express the "transformation of wave functions when the frame is rotated." Note that if $(g,\mathbf{v})$ is any Euclidean motion, (68) tells us that we must first transform ψ into the function $\mathbf{x} \rightarrow \psi(g^{-1}\mathbf{x} - g^{-1}\mathbf{v})$, and then multiply the result by the matrix $D^j(g^*)$, where g^* lies above g. There is of course an ambiguity in the choice of g^*; but this only changes $D^j(g^*)$ into $\pm D^j(g^*)$ and the resulting lattice automorphism depends only on g. For integral values of j, there is no need to go to G^*, as D^j defines already a representation of G_0 itself.

From (68) we can obtain the expressions for the various momenta corresponding to the elements of the Lie algebra of G. If $\mathbf{v} \in \mathbf{R}^3$, and we consider the translations $\mathbf{x} \rightarrow \mathbf{x} + t\mathbf{v}$ of $\mathbf{R}^3$, we obtain, for the corresponding linear momentum $B_{\mathbf{v}}$, the expression

$$B_{\mathbf{v}} : \psi \rightarrow -i\partial(\mathbf{v})\psi, \tag{69}$$

where $\partial(\mathbf{v})$ operates componentwise. This is exactly the same as in (32). On the other hand, we now have in addition the *angular momenta*, which correspond to the various one-parameter subgroups of G_0^*. Let X_1, X_2, and X_3 be the elements of the Lie algebra of G_0^* which correspond to the one-parameter groups g_1, g_2, and g_3, respectively. If M_1, M_2, and M_3 are the corresponding momentum operators, we have, from (62), (63), and (68), for $r = 1, 2, 3$,

$$M_r\psi = i\{\partial_r\psi + \dot{D}^j(X_r)\psi\}, \tag{70}$$

where

$$\begin{aligned} \partial_1 &= -x_3 \frac{\partial}{\partial x_2} + x_2 \frac{\partial}{\partial x_3}, \\ \partial_2 &= x_3 \frac{\partial}{\partial x_1} - x_1 \frac{\partial}{\partial x_3}, \\ \partial_3 &= -x_2 \frac{\partial}{\partial x_1} + x_1 \frac{\partial}{\partial x_2}, \end{aligned} \tag{71}$$

and

$$\dot{D}^j(X_r) = \left(\frac{d}{dt} D^j(g_r^*(t))\right)_{t=0}. \tag{72}$$

It is usual to write M_r^0 for $i\partial_r$ and M_r^s for the operator $\psi \rightarrow i\dot{D}^j(X_r)\psi$. M_r is called the *angular momentum around the x_r-axis*. The operators M_r^0 and M_r^s are actually defined by the relations

$$\begin{aligned} \exp(-itM_r^0) &= A_t^r, \\ \exp(-itM_r^s) &= B_t^r, \end{aligned}$$

where $t \to A_t^r$ is the one-parameter group defined by $(A_t^r\psi)\cdot(\mathbf{x}) = \psi(\delta(g_r^*(t))^{-1}\mathbf{x})$, and $t \to B_t^r$ is the one-parameter group defined by $(B_t^r\psi)(\mathbf{x}) = D^j(g_r^*(t))\psi(\mathbf{x})$. Since A_t^r and B_t^k commute, so do M_r^0 and M_k^s:

$$[M_r^0, M_k^s] = 0 \qquad (r, k = 1, 2, 3). \tag{73}$$

M_r^0 is called the *orbital angular momentum* around the x_r-axis, and M_r^s is called the *spin angular momentum* around the x_r-axis. By lemma 8.2, M_r^0 is essentially self-adjoint on $C_c^\infty(\mathbf{R}^3, W)$; M_k^s is bounded. Moreover, M_r is also essentially self-adjoint on $C_c^\infty(\mathbf{R}^3, W)$.

We note that the group G_0^*, as well as its one-parameter subgroups, is *compact.* Hence for any element X of the Lie algebra of G_0^*, the operator B_X (defined by $\exp(-itB_X) = V_{(\exp tX)}$ will have *a discrete spectrum.* From the formulas (62) it follows at once that the angular momentum observables M_r have only the half integral numbers $0, \pm\frac{1}{2}, \pm 1, \pm\frac{3}{2}, \cdots$ as their possible values. This was one of the earliest facts recognized in quantum mechanics and emphasized by Niels Bohr (cf. Weyl [1], pp. 185-191).

The formulas (54) show that insofar as the translations of M are concerned, we have the usual system of imprimitivity for the translation group. However, if $j \neq 0$, this restricted system is no longer irreducible. Therefore there will be observables which will be simultaneously measurable with respect to the linear momenta and the configuration observables. The operators M_1^s, M_2^s, and M_3^s correspond to such observables as can be directly verified from the equation

$$M_r^s\psi = i\dot{D}^j(X_r)\psi.$$

These are usually called the *spin observables.* Note that each M_r^s acts only on the components of ψ, and not on the argument $\mathbf{x}$ of ψ. The possible values of M_r^s are just $2j+1$ in number; these are $j, j-1, j-2, \cdots, -j$. The parameter j, which enters so basically in the description of the system, is called the *spin of the system.*

The case $j = \frac{1}{2}$ is of special interest. In this case, the ψ's map $\mathbf{R}^3$ into $\mathbf{C}^2$, so that each ψ can be written as $\begin{pmatrix}\psi_1\\ \psi_2\end{pmatrix}$, where ψ_1 and ψ_2 are square integrable on $\mathbf{R}^3$. A brief calculation then shows that the operators M_r^s are described by

$$M_r^s : \psi \to -\tfrac{1}{2}\sigma_r\psi, \tag{74}$$

where σ_1, σ_2, and σ_3 are described by (55). The observables M_r^s have only the two possible values $\pm\frac{1}{2}$.

With this we have completed what might be called the *kinematical description* of covariant systems with a Euclidean configuration space. We have obtained the usual expressions for the familiar observables and shown that the introduction of the spin observables is an inevitable consequence of the *geometry* of the configuration space. This seems to be the mathematical meaning of spin.

6. PARTICLES

We shall now use the analysis given in Sections 1 through 5 to develop some of the formulas and expressions which are at the heart of non-relativistic quantum mechanics, namely those dealing with the mechanics of one particle. We shall, however, not go into many details. In the next chapter we shall make a deeper examination of the relativistic nature of a particle. Our present remarks are thus intended to be brief and serve only the purpose of a formal description of a few standard facts of quantum mechanics.

We shall first examine the usual description of a "free" particle. Clearly one would like to associate with a free particle a configuration space M and therefore give meaning to statements about the *position* of the particle. At the same time, the basic nature of such a system requires a covariance with respect to the whole group of motions of M. We assume that M is a three-dimensional Euclidean *affine* space. Then the results of Section 5 may be applied to yield the result that these systems are labeled by the spin parameter j. To specify the system completely we must determine also its Hamiltonian. Now the Hamiltonian H cannot be an arbitrary self-adjoint operator since one would like our system to possess the usual rules of *conservation of momenta.* According to our principles (cf. Section 1), the requirement that the momenta be integrals of motion leads at once to the condition that the dynamical operators $\exp(-itH)$ commute with the operators V_{g^*} which correspond to the various elements g^* of G^* (cf. (13), (14)).

Let us now determine the most general self-adjoint operator H such that $\exp(-itH)$ commutes with all the operators of the representation V. We consider the case when the spin is 0. We choose a point $x_0 \in M$ and a Euclidean coordinate system with x_0 as origin. We can then identify M with $\mathbf{R}^3$ and the Hilbert space of the system with $\mathscr{L}^2(\mathbf{R}^3)$. If $(g^*,\mathbf{v}) \in G^*$ and $\psi \in \mathscr{L}^2(\mathbf{R}^3)$,

$$(V_{(g^*,\mathbf{v})}\psi)(\mathbf{x}) = \psi(\delta(g^*)^{-1}(\mathbf{x}-\mathbf{v})) \qquad (\mathbf{x} \in \mathbf{R}^3).$$

Let $\mathscr{F} : \psi \to \hat{\psi}$ be the Fourier-Plancherel transformation of $\mathscr{L}^2(\mathbf{R}^3)$ onto itself (cf. (41)). Then a straightforward calculation shows that

$$(\mathscr{F} V_{(g^*,\mathbf{v})} \mathscr{F}^{-1}\varphi)(\mathbf{x}) = \exp[-i\langle \mathbf{x},\mathbf{v}\rangle]\cdot\varphi(\delta(g^*)^{-1}\mathbf{x}). \tag{75}$$

Let $K = \mathscr{F} H \mathscr{F}^{-1}$. Then the fact that K commutes with all the operators $\mathscr{F} V_{(1,\mathbf{v})} \mathscr{F}^{-1}$ implies that K is the operator of multiplication by some Borel function k (possibly unbounded). If we now use the condition that K commutes with all the operators $\mathscr{F} V_{(g^*,0)} \mathscr{F}^{-1}$, we can conclude that for each rotation g, $k(\mathbf{x}) = k(g^{-1}\mathbf{x})$ for almost all $\mathbf{x}$. From this it follows quite simply that k must coincide almost everywhere with a Borel function of $x_1{}^2 + x_2{}^2 + x_3{}^2$. We may therefore conclude at once that H must be a Borel function, in the sense of the functional calculus of spectral theory, of the *Laplacian*, ∇:

$$\nabla = \frac{\partial^2}{\partial x_1{}^2} + \frac{\partial^2}{\partial x_2{}^2} + \frac{\partial^2}{\partial x_3{}^2}. \tag{76}$$

A similar argument shows that, conversely, for any Borel function $k_1(\nabla)$, the operators $\exp[-itk_1(\nabla)]$ commute with all the operators of the representation V, so that in the system with Hamiltonian $k_1(\nabla)$, all the linear and angular momenta will be conserved.

Covariance with respect to the group of motions of M cannot therefore pinpoint the Hamiltonian H of a free particle other than up to a function of ∇. The usual form of H can be obtained only after we introduce the notion of *relativity*. If we demand that the system be covariant (in a suitable sense) under the full space-time Galilean group, then one can show that the only possible Hamiltonians are of the form

$$H = -\frac{1}{2m} \nabla, \tag{77}$$

where m is a real number; we choose $m > 0$ so that the spectrum of H is bounded below (positivity of energy). We shall not go through the calculations here; we shall prove (77) in the next chapter, where we take care of the case of arbitrary spin also.

The formula (77) is truly remarkable; for, if we notice that the operator $p_r = -i(\partial/\partial x_r)$ corresponds to the linear momentum along the x_r-axis, then (77) can be rewritten as

$$H = \frac{1}{2m} (p_1{}^2 + p_2{}^2 + p_3{}^2). \tag{78}$$

Formula (78) is also true if we interpret both sides classically; H is then the well known Hamiltonian for a free particle of mass m. Thus equation (78) further deepens the formal analogy between classical and quantum mechanics.

If we are dealing not with a free particle but one moving in a potential field, then it is natural to assume that the expressions for the momentum and position observables retain their validity. But the expression for energy will now have to be different. It is natural to retain (78) for the *kinetic energy* of the particle and add to it a term involving the *potential energy*. Now the potential energy must be regarded as a function of the position of the particle only and hence must be represented by an operator which is a function of q_1, q_2, and q_3. We thus arrive at the expression of the Hamiltonian for a particle moving in a potential field:

$$H_S = \frac{1}{2m}(p_1{}^2+p_2{}^2+p_3{}^2)+S(q_1,q_2,q_3), \tag{79}$$

where $S(q_1,q_2,q_3)$ is the operator of multiplication by the function S on $\mathbf{R}^3$; the operator S represents the potential energy of the particle. Strictly speaking, one must be rather careful about (79), since we are adding to the operator $(1/2m)(p_1{}^2+p_2{}^2+p_3{}^2)$ another self-adjoint operator of possibly unbounded character. However, in a large number of physically interesting cases there is no ambiguity about the definition of H_S as a self-adjoint operator. The formula (79) would then complete the description of the one-particle systems.

Consider now a special case when S is a C^∞ function and of a sufficiently simple type so that H_S is essentially self-adjoint on the space of C^∞ functions with compact supports. If we now use the Heisenberg picture and compute the time derivatives of the p_r's and q_r's we find, from (79) and (14), on $C_c{}^\infty(R^3)$,

$$\begin{aligned} \dot{p}_r &= -\partial H_S/\partial q_r, \\ \dot{q}_r &= \partial H_S/\partial p_r, \end{aligned} \tag{80}$$

where $\partial H_S/\partial q_r$ is the operator of multiplication by $\partial S/\partial q_r$, and $\partial H_S/\partial p_r$ is the operator $(1/m)p_r$. These are the analogues of the classical Hamiltonian equations in their operatorial form. They formed the point of departure for Heisenberg's classic approach to the problems of the atomic systems. On the other hand, (79) was the starting point of Schrödinger's investigations. These two approaches, together with the statistical interpretation of Max Born (cf. (39)), were the beginning of the epoch which led to the modern revolution in physics (cf. Weyl [1]).

We conclude with two remarks bearing on our special assumptions on the configuration space. Our main assumption is that M is a Euclidean affine space. This in fact was decisive since it determined the structure of its group of motions and cleared the way for the application of the main theorems dealing with systems of imprimitivity. However, there are other interesting classes of spaces on which the motions act transitively; for example, the Riemannian manifolds with *constant* negative curvature (cf. Helgason [1]). In this case, it would be of interest to carry out an analysis analogous to what we have done, at least when the curvature is small. In particular, it would be of interest to compute the spectra of various important Hamiltonians as functions of the curvature ε, for small $\varepsilon<0$.

Our second remark concerns the existence of configuration space. This led to our being able to calculate, for any state of our system, the probability that a particle is found in a specified region of the space M. However, there are interesting physical arguments which prohibit, for certain particles like the photon, the possibility of experimentally verifying whether they are in a specified region of space. Such particles cannot be described, even kinematically, by the considerations of this chapter. We shall examine these questions from a different and more unified point of view in the next chapter.

NOTES ON CHAPTER VIII

1. *Quantization.* Roughly speaking, Quantization refers to a process by which one establishes a "correspondence" between certain classical and quantum systems. The mathematical requirements of this correspondence are, however, capable of being formulated in many ways and so there are several approaches to quantization. Historically it was Niels Bohr who first understood the importance, and even the necessity of constructing such correspondences. His celebrated "correspondence principle" was based on his discovery that for electron orbits with "large" quantum numbers, his new frequency energy relation approximated quite closely the classical relation; here large means large in relation to $\hbar$. Thus it became possible to view Quantum Mechanics as a generalization of classical mechanics whose results were approximated by those furnished by classical theories when $\hbar$, now treated as a parameter, tends to 0. With the discovery of Quantum Mechanics by Heisenberg, Schrödinger, and Dirac, the relationship between classical and quantum theories became clearer. In the Schrödinger representation, the classical state space is the phase space $\mathbf{R}^{2n}=\mathbf{R}^n\times\mathbf{R}^n$ while the quantum Hilbert space is $L^2(\mathbf{R}^n)$, and one "associates" to classical observables which are functions of $q_1,\dots,q_n$, $p_1,\dots,p_n$, the quantum observables obtained by the correspondence

$$q_j\to Q=\text{multiplication by } q_j,$$

$$p_j\to -i\hbar\frac{\partial}{\partial q_j}=P_j.$$

However, since the P_j and Q_j do not commute, such correspondences are not precisely defined even at the formal level, and so one needs additional prescriptions. The first rigorous quantization appears to be that of Hermann Weyl [1] (pp. 272–280) who associated to the classical function

$$a=a(q_1,\dots,q_n,p_1,\dots,p_n)$$

the operator A where

$$A = \int \hat{a}(u_1, \ldots, u_n, v_1, \ldots, v_n) \times \exp\{i\hbar^{-1}(u_1P_1 + \ldots + u_nP_n + v_1Q_1 + \ldots + v_nQ_n)\}\, du_1 \ldots du_n\, dv_1 \ldots dv_n,$$

where $\hat{a}$ is the inverse Fourier transform of a and the exponential is interpreted via the unitary representation of the Heisenberg group. Since then this theme has been examined repeatedly and from many other points of view. It is natural to replace the classical phase space $\mathbf{R}^{2n}$ by a symplectic manifold of dimension $2n$, or at least by the cotangent bundle of a "configuration manifold" of dimension n. The questions involve the differential geometry and cohomology of these manifolds as well as the analytical aspects of symplectic geometry and Fourier transforms. The following is a partial list of sources.

M. Flato and D. Sternheimer, in *Harmonic Analysis and representations of semisimple Lie groups*, edited by J. A. Wolf, M. Cahen, and M. de Wilde, Reidel, Dordrecht, 1980, pp. 385–448.

V. Guillemin and S. Sternberg, *Geometric Asymptotics*, Mathematical Surveys, No. 14, Amer. Math. Soc., Providence, R.I., 1977.

B. Kostant, in *Lectures in Modern Analysis and Applications*, III. Lecture Notes in Mathematics, No. 170, Springer-Verlag, Berlin, 1978, pp. 87–208.

J. Leray, *Lagrangian Analysis and Quantum Mechanics*, M.I.T Press, Cambridge, Mass., 1981.

V. P. Maslov, *Theorie des perturbations et methodes asymptotiques, suive de deux notes complementaires de V. I. Arnold et V. C. Buslaev*, Paris, Dunod, 1972 (tr. by J. Lascoux and R. Senor).

J. E. Moyal, Proc. Cambridge Phil. Soc., **45** (1949), pp. 99–124.

J. M. Souriau, *Construction explicite de l'indice de Maslov Applications*, Fourth International Colloquium on Group Theoretical Methods in Physics, University of Nijmegen, 1975.

2. *Second Quantization.* This is a framework for treating systems of identical particles from the quantum mechanical point of view and studying their behavior when the number of particles is large. It goes back to P. A. M. Dirac's famous paper on radiation theory (Proc. Roy. Soc. London, A, **114** (1927), pp. 243–265) and is fundamental in quantum statistical physics and quantum field theory. If $\mathscr{K}$ is the Hilbert space associated with a single particle, then second quantization associates to $\mathscr{K}$ a Hilbert space $\mathscr{H}$ in an invariant manner; $\mathscr{H}$ will be the space corresponding to an unlimited number of particles identical with the original one and will have additional structure determined by the existence of certain operators. The key features present in $\mathscr{H}$ may be described as follows:

(a) *The operator N = the number of particles.* There is a self-adjoint operator N with pure discrete spectrum and eigenvalues $0, 1, 2, \ldots$; if $\mathscr{H}_n$ is the eigenspace for the eigenvalue n, states corresponding to the rays in $\mathscr{H}_n$ are those in which the system has exactly n particles. Also

$$\mathscr{H}_1 = \mathscr{K}.$$

The vectors of $\mathscr{H}_n$ may be denoted by $|n\rangle$; thus $N|n\rangle = n|n\rangle$.

(b) *Vacuum state.* The space $\mathscr{H}_0$ is one dimensional,

$$\dim(\mathscr{H}_0) = 1.$$

The state corresponding to this space is called the *vacuum.*

(c) *Occupation numbers: The operators $N(\varphi)$.* If $(\varphi_i)_{i\geqslant 1}$ is an orthonormal basis for $\mathscr{K}$ and $(n_i)_{i\geqslant 1}$ are integers $\geqslant 0$, there are unit vectors, denoted by $|(n_i)\rangle$, with the property that for any i, n_i particles are in the state φ_i. The n_i are the so-called occupation numbers. This is formulated by requiring the existence of self-adjoint operators $N(\varphi)$ with pure discrete spectrum consisting of $0, 1, \ldots$, the number of particles in state φ. The $N(\varphi_i)$ commute among themselves, $N(\varphi_j)|(n_i)\rangle = n_j|(n_i)\rangle$, and for any $n \geqslant 0$, the $|(n_i)\rangle$ with $n_1 + \ldots = n$ form an orthonormal basis for $\mathscr{H}_n$.

(d) *Creation and Annihilation operators.* For each vector φ in the one-particle space $\mathscr{K}$ there is an operator $A(\varphi)$ in $\mathscr{H}$, which "annihilates" one particle in state φ, and its adjoint $A(\varphi)^\dagger$, which "creates" one particle in state φ. These operators could be unbounded. We require

$$A(\varphi)^\dagger A(\varphi) = N(\varphi).$$

If (φ_i) is an orthonormal basis of $\mathscr{K}$, it is clear from the occupation number interpretation that $A(\varphi_l)^\dagger A(\varphi_m)^\dagger$ and $A(\varphi_m)^\dagger A(\varphi_l)^\dagger$ should both take $|(n_i)\rangle$ to vectors which are proportional. The simplest way to ensure this is to require that $A(\varphi_l)^\dagger A(\varphi_m)^\dagger = cA(\varphi_m)^\dagger A(\varphi_l)^\dagger$; then $c = \pm 1$, and one is thus led to distinguish between two types of systems: *Bose–Einstein* (B–E) and *Fermi–Dirac* (F–D). The requirements are (with $[A,B] = AB - BA$, $\{A,B\} = AB + BA$)

(B–E)
(i) $A(\varphi)$ antilinear in φ;
(ii) $[A(\varphi), A(\psi)] = 0$;
(iii) $[A(\varphi), A(\psi)^\dagger] = \langle\psi,\varphi\rangle 1$.

(F–D)
(i) $A(\varphi)$ antilinear in φ;
(ii) $\{A(\varphi), A(\psi)\} = 0$;
(iii) $\{A(\varphi), A(\psi)^\dagger\} = \langle\psi,\varphi\rangle 1$.

(In this brief discussion we do not insist on domain considerations in interpreting these rules.) The corresponding particles are respectively known as

Bosons and *Fermions.* It is a fundamental empirical fact that particles with *integral* spin are Bosons while those with *half-integral* spin are Fermions.

(e) *Cyclicity of vacuum.* It is natural to assume that

$$A(\varphi)|0\rangle = 0.$$

The cyclicity assumption means that the vectors

$$A(\varphi_1)^\dagger A(\varphi_2)^\dagger \ldots A(\varphi_m)^\dagger|0\rangle \qquad (\varphi_i \in \mathscr{K})$$

span a dense subspace of $\mathscr{H}$.

(f) *Covariance.* There is a strongly continuous unitary representation Γ of the unitary group of $\mathscr{K}$ in $\mathscr{H}$, commuting with N. Thus $\Gamma(U)\mathscr{H}_n = \mathscr{H}_n$ for all n, $U \in$ unitary group of $\mathscr{K}$.

It is not difficult to show that once the type of the particle is specified (that is, a Boson or Fermion), then there is essentially one $\mathscr{H}$ associated to $\mathscr{K}$. $\mathscr{H}$ is called the *Fock space*; it was first discovered by V. Fock. We give a brief description of it.

Bosons. $\mathscr{H} = S(\mathscr{K})$, the *symmetric algebra* over $\mathscr{K}$. This is defined essentially in the same way as in the case when $\dim(\mathscr{K}) < \infty$. More precisely, let $T(\mathscr{K}) = T$ be the covariant tensor algebra of $\mathscr{K}$, i.e.,

$$T = \mathbf{C}1 \oplus \mathscr{K} \oplus (\mathscr{K} \otimes \mathscr{K}) \oplus \ldots = \bigoplus_{m \geqslant 0} (\mathscr{K}^{\otimes m}),$$

where we take the infinite Hilbert space direct sum. Let J_s be the *closed* linear span of all tensors of the form

$$u \otimes (x \otimes y - y \otimes x) \otimes v,$$

where $x, y \in \mathscr{K}$, $u, v \in \sum_{m \geqslant 0} \mathscr{K}^{\otimes m} = T^0$. Then

$$S \underset{\text{dfn}}{=} T/J_s.$$

As in the finite dimensional case, we have a linear Hilbert space isomorphism

$$S \simeq T^{\text{symm}},$$

where T^{symm} is the subspace of symmetric tensors:

$$T^{\text{symm}} = \bigoplus_{m \geqslant 0} (\mathscr{K}^{\otimes m})^{\text{symm}}.$$

S is of course graded, the grading being inherited from T:

$$S = \bigoplus_{m \geqslant 0} S_m, \qquad \mathscr{H}_m = S_m;$$

and the *algebraic sum*

$$S^0 \underset{\text{dfn}}{=} \sum_{m \geqslant 0} S_m$$

is an *algebra*. The operators $A(\varphi)$ and $A(\varphi)^\dagger$ are unbounded, they are first defined on S^0 and then extended by formation of closures. On S^0 they are defined by

$$A(\varphi)^\dagger : u \mapsto \sqrt{m+1}\,\varphi u \qquad (u \in S_m),$$

$$A(\varphi) : u \mapsto \frac{1}{\sqrt{m}}\,\partial(\varphi)u \qquad (u \in S_m),$$

where $\partial(\varphi)$ is "differentiation" along φ, defined as the unique derivation of S^0 such that

$$\partial(\varphi)\,\varphi' = \langle \varphi', \varphi \rangle\, 1.$$

The verification that these definitions lead to a B–E system is straightforward. If U is a unitary operator of $\mathscr{K}$, it lifts naturally to T, and hence induces a unitary automorphism $\Gamma(U)$ of S which leaves S^0 invariant and is an algebra automorphism of it. Of course $|0\rangle = 1$.

Fermions. The relations

$$A(\varphi)^\dagger A(\varphi) + A(\varphi) A(\varphi)^\dagger = 1$$

show that the $A(\varphi)$ and $A(\varphi)^\dagger$ are *bounded operators*. For the standard model we choose $\mathscr{H} = \Lambda$, the *exterior algebra of* $\mathscr{K}$. It is defined as T/J^a where J_a is the closed linear span containing all tensors of the form

$$u \otimes (x \otimes y + y \otimes x) \otimes v,$$

where $x, y \in \mathscr{K}$, $u, v \in T^0$. Λ is graded in the obvious way; $\Lambda_m = \mathscr{H}_m$ is the image of $\mathscr{K}^{\otimes m}$ under $T \to \Lambda$. If we write

$$\Lambda^0 = \sum_{m \geqslant 0} \Lambda_m \qquad \text{(algebraic sum)},$$

Λ^0 is an algebra. As usual $\wedge$ denotes multiplication in Λ. The creation and annihilation operators are defined by

$$A(\varphi)^\dagger : u \mapsto \sqrt{m+1}\,\varphi \wedge u \qquad (u \in \Lambda_m),$$

$$A(\varphi) : u \mapsto \frac{1}{\sqrt{m}}\,\partial(\varphi)u \qquad (u \in \Lambda_m),$$

where $\partial(\varphi)$ is the unique endomorphism of Λ^0 such that

(i) $\partial(\varphi)1 = 0,$

(ii) $\partial(\varphi)\varphi' = \langle \varphi', \varphi \rangle 1,$

(iii) $\partial(\varphi)(u_m \wedge u_n) = (\partial(\varphi)u_m) \wedge u_n + (-1)^m u_m \wedge \partial(\varphi)u_n$
$(u_m \in \Lambda_m,\ u_n \in \Lambda_n).$

If U is a unitary operator of $\mathscr{K}$, $\Gamma(U)$ is the corresponding naturally induced unitary operator of Λ; $\Gamma(U)$ is an algebra automorphism of Λ^0. Finally $|0\rangle = 1$.

In the Fock representation for a B–E system choose a real form $\mathscr{K}_{\mathbf{R}}$ of $\mathscr{K}$ and write

$$P(\varphi) = \frac{A(\varphi) + A(\varphi)^{\dagger}}{\sqrt{2}}, \qquad Q(\varphi) = \frac{A(\varphi)^{\dagger} - A(\varphi)}{i\sqrt{2}}.$$

Then we have

$$[P(\varphi), P(\psi)] = 0, \qquad [Q(\varphi), Q(\psi)] = 0 \qquad [Q(\varphi), P(\psi)] = i\langle\varphi, \psi\rangle.$$

These are analogues of the Heisenberg commutation rules, and reduce to them when $\dim(\mathscr{K}) < \infty$. The uniqueness of the irreducible representation of the Heisenberg commutation rules (in the "Weyl" or "integrated" form) when $\dim(\mathscr{K}) < \infty$ makes clear why the B–E system is unique. We can in fact use the Schrödinger representation to construct another model for the Fock representation. Let $\mathscr{H} = L^2(\mathscr{K}_{\mathbf{R}})$ (with respect to the Lebesgue measure) and let

$$P(\varphi) = \frac{1}{i}\partial(\varphi), \qquad Q(\varphi) = \langle\varphi, .\rangle \qquad \text{(multiplication by } \langle\varphi, .\rangle).$$

Then it is enough to take

$$A(\varphi) = \frac{P(\varphi) - iQ(\varphi)}{\sqrt{2}}, \qquad A(\varphi)^{\dagger} = \frac{P(\varphi) + iQ(\varphi)}{\sqrt{2}}.$$

The vacuum vector is the function

$$\exp(-\tfrac{1}{2}\|\cdot\|^2).$$

This type of model can be constructed even when $\dim(\mathscr{K}) = \infty$. It is, however, technically more complicated because there is no translation invariant (or translation quasi invariant) measure on $\mathscr{K}_{\mathbf{R}}$ when $\dim(\mathscr{K}) = \infty$, and one has to use a more sophisticated type of integration theory, namely integration with respect to the isotropic Gaussian weak distribution dg on $\mathscr{K}_{\mathbf{R}}$ (cf. I. E. Segal in *Les Problèmes mathématiques de la théorie quantique des champs*, CNRS, Paris, 1959, pp. 57–103).

The identification $L^2(\mathscr{K}_{\mathbf{R}}, dg) \approx \mathscr{H}$ allows us to "quantize" any classical dynamical system whose configuration space is $\mathscr{K}_{\mathbf{R}}$ and whose dynamics is *linear and unitary on* $\mathscr{K}$. If (U_t) is the classical dynamical group, $(\Gamma(U_t))$ is the quantized dynamical group; moreover, the grading $\mathscr{H} = \bigoplus_n \mathscr{H}_n$ leads to a particle interpretation. This was how Dirac quantized the electromagnetic field and constructed the quantum theory of radiation, namely interaction of light with matter.

For the Fermion case, we keep the real form $\mathscr{K}_{\mathbf{R}}$ as before but now define

$$P(\varphi) = A(\varphi) + A(\varphi)^{\dagger}, \qquad iQ(\varphi) = A(\varphi)^{\dagger} - A(\varphi).$$

Then P and Q are linear maps of $\mathscr{K}_{\mathbf{R}}$ into the space of bounded self-adjoint operators on $\mathscr{H}$ satisfying

$$P(\varphi)^2 = Q(\varphi)^2 = \langle\varphi,\varphi\rangle 1, \qquad \{P(\varphi),Q(\psi)\} = 0.$$

This suggests the introduction of the *Clifford algebra*, Cliff($\mathscr{L}$) associated to a real Hilbert space $\mathscr{L}$, which is a complex algebra with unit 1, containing $\mathscr{L}$ and generated by it, such that

$$u^2 = \langle u,u\rangle 1 \qquad u \in \mathscr{L}.$$

Cliff($\mathscr{L}$) exists and is essentially uniquely determined by $\mathscr{L}$. If we put

$$R(\varphi + i\psi) = Q(\varphi) + P(\psi), \qquad \varphi, \psi \in \mathscr{K}_{\mathbf{R}},$$

and write $\underline{\mathscr{K}}$ for the real Hilbert space underlying $\mathscr{K}$,

$$(\langle u,v\rangle_{\underline{\mathscr{K}}} = \mathrm{Re}\langle u,v\rangle)$$

then

$$(P,Q) \leftrightarrow R$$

is a bijection between F–D systems based on $\mathscr{K}_{\mathbf{R}}$ and representations R of Cliff ($\underline{\mathscr{K}}$) such that $R(u)$ is bounded self-adjoint for all $u \in \underline{\mathscr{K}}$.

For a more systematic study of the framework of second quantization the reader may refer to F. A. Berezin, *The Method of Second Quantization*, Academic Press, New York, 1966, and J. M. Cook, *The Mathematics of Second Quantization*, Trans. Amer. Math. Soc., **74** (1953), pp. 222–245. In the original article of Dirac referred to earlier, he used this framework to study the interaction of an atom with the electromagnetic field and obtained the known formulas for probabilities of transitions between the energy levels. For a detailed treatment see Dirac, loc. cit., and von Neumann [1] (pp. 254–294). Dirac's paper inaugurated the beginning of an entire era in which a major goal was to erect a theory of quantized fields obeying the principle of special relativity. To get a picture of what has been done, as well as what remains to be done, we refer the reader to the following books and to the articles referred to in them:

J. Schwinger (Ed.), *Selected Papers on Quantum Electrodynamics*, Dover, New York, 1958.

S. Schweber, *An Introduction to Relativisitic Quantum Field Theory*, Row, Peterson and Co., New York, 1961.

J. M. Jauch and F. Rohrlich, *The Theory of Photons and Electrons*, Springer-Verlag, Berlin, 1980.

N. N. Bogoliubov and D. V. Shirkov, *Introduction to the Theory of Quantized Fields*, Wiley, New York, 1980.

J. Glimm and A. Jaffe, *Quantum Physics: A Functional Integral Point of View*, Springer-Verlag, Berlin, 1981.

CHAPTER IX

RELATIVISTIC FREE PARTICLES

1. RELATIVISTIC INVARIANCE

The fundamental assumption in Newtonian mechanics is the absolute nature of space and time. That our empirical knowledge of the physical world points up to the contrary was one of the cornerstones of Einstein's great critique of space and time. The analysis of Einstein and Lorentz of the empirical and mathematical nature of the physical phenomena established that only *space-time*, as a four-dimensional manifold, has an invariant physical significance, and that its familiar splitting into space and time is essentially dependent on the observer. Each observer was thus seen to describe only the points of the space-time manifold in his vicinity by four coordinates; the formulas describing the transformation of coordinates of the same space-time point by two different observers involved smooth functions. In other words, space-time became an abstract C^∞ manifold and each observer was seen to describe a neighborhood of this manifold in terms of local coordinates. Only statements which retain their meaning in every coordinate system were asserted to have any physical significance, i.e., the *global differential geometric* statements. Einstein's theory of gravitation is, for example, a formulation of the theory of gravitational phenomena involving only intrinsic geometric objects like connections and tensor fields.

In problems dealing with the quantum aspects of atomic systems, it is usual to assume that gravitational fields have no influence on the physical phenomena (cf. Dirac [1], p. 253). As a consequence of this assumption, the group of transformations which connect the different observers is drastically reduced. We shall make the assumption that the observers are what are known as inertial, and that the relevant group of transformations is the so-called group of special relativity. We shall now proceed to make a few explanatory remarks on these concepts.

In special relativity it is assumed that each observer is able to establish coordinates for the whole of the space-time manifold and, moreover, that he uses a rather special coordinate system to achieve this. One makes the assumption that the notion of a *classical free particle*, not acted upon by external forces, is meaningful for the observer, and that the trajectories of such a particle are described by linear functions of the time coordinate in his coordinate system. It is customary to call such coordinate systems and observers *inertial*. It may be noted that an inertial coordinate system is simply one in which the first law of Newton is valid—every particle is at rest or continues in its state of uniform motion in a straight line, unless it is acted upon by external forces to do otherwise. Given two inertial coordinate systems, say O, O', each of which maps the space-time manifold in a one-one fashion onto $\mathbf{R}^4$, a correspondence $T_{O,O'}$ is established between the points of $\mathbf{R}^4$; (x_0,x_1,x_2,x_3) and (x_0',x_1',x_2',x_3') correspond under $T_{O,O'}$ if and only if they represent the same space-time point, in O and O', respectively.†

What can one say about the mapping $T_{O,O'}$,

$$T_{O,O'} : (x_0,x_1,x_2,x_3) \to (x_0',x_1',x_2',x_3')? \tag{1}$$

It is natural to expect $T_{O,O'}$ to be an *inhomogeneous linear* transformation; for, as both coordinate systems are inertial, the trajectories of a free particle must be represented by straight lines in *both* systems, i.e., $T_{O,O'}$ must map straight lines into straight lines. Hence it must preserve coplanarity and parallelism. Consequently, if we pass (through the use of homogeneous coordinates) to the associated *projective geometry* of R^5, $T_{O,O'}$ will induce an automorphism of this projective geometry and is therefore described in $\mathbf{R}^5$ by a unique *linear* automorphism (cf. the results of Chapter II); back in $\mathbf{R}^4$, this would lead to the affine nature of $T_{O,O'}$. It is roughly along these lines that one is led to the basic hypothesis of special relativity: *space-time is an affine space‡; each inertial observer describes it with an affine coordinate system, and the transformations connecting the coordinates of the same point in two different inertial systems are inhomogeneous linear:*

$$x_i' = \sum_{j=0}^{3} a_{ij}x_j + u_i. \tag{2}$$

For three inertial observers O, O', and O'', we have:

$$T_{O,O''} = T_{O',O''}T_{O,O'}, \tag{3}$$

$$T_{O',O} = T^{-1}_{O,O'}. \tag{4}$$

Clearly the set of all $T_{O,O'}$ forms a group. This group is called the *space-time relativity group*. It has been obtained explicitly as a group of inhomogeneous linear automorphisms of $\mathbf{R}^4$. In the abstract affine space-time it gives rise to a group of affine automorphisms.

---

† x_0 and x_0' are the time coordinates: x_j, x_j' $(j=1, 2, 3)$ the space coordinates.

‡ Cf. Chapter VIII, Section 4.

However, not every inhomogeneous map of the form (2) can arise as a transformation between inertial coordinate systems. Further examination of the empirical nature of space-time phenomena yields certain restrictions on the $T_{O,O'}$. Before Einstein and Lorentz, it used to be assumed implicitly that the notion of *simultaneity* of two events was an invariant one, i.e., if O, O' are two inertial coordinate systems, if (x_0,x_1,x_2,x_3) and (y_0,y_1,y_2,y_3) are the points of R^4 which correspond under O to two space-time points, and $(x_0',x_1'x_2',x_3')$ and (y_0',y_1',y_2',y_3') represent the same pair of points under O', then $x_0'=y_0'$ if and only if $x_0=y_0$. It follows from this that in the formula (2) for $T_{O,O'}$, we must have

$$x_0' = a_{00}x_0+u_0. \tag{5}$$

If we assume that the unit of time measurement and the sense of time propagation is the same in both systems, we must have $a_{00}=1$, i.e.,

$$x_0' = x_0+u_0. \tag{6}$$

Furthermore it was assumed that the *spatial distance* of pairs of space-time points with the *same* time coordinate also remained unchanged, i.e.,

$$\sum_{i=1}^{3} (x_i-y_i)^2 = \sum_{i=1}^{3} (x_i'-y_i')^2. \tag{7}$$

From this it follows that the matrix of order 3×3 in (2),

$$W = (a_{ij})_{1\leq i,j\leq 3} \tag{8}$$

is orthogonal. Consequently, if we write $\mathbf{x}$ for the column vector with components x_1, x_2, and x_3,

$$\begin{aligned} \mathbf{x}' &= W\mathbf{x}+x_0\mathbf{v}+\mathbf{u}, \\ x_0' &= x_0+u_0, \end{aligned} \tag{9}$$

where W is an orthogonal 3×3 matrix, $\mathbf{v},\mathbf{u}\in\mathbf{R}^3$, and $u_0\in\mathbf{R}^1$. The collection of all transformations $(x_0,\mathbf{x})\rightarrow(x_0',\mathbf{x}')$ of the form (9) is easily verified to be a group.

We now assert that every transformation of the form (9) is a possible transformation between two inertial coordinate systems. Since the possible transformations constitute a group, it is enough to exhibit a class of transformations which actually arises in practice and which also generates the group of all transformations of the form (9). If we put in (9) $u_0=0$, $\mathbf{u}=\mathbf{v}=0$, we obtain a transformation which connects two coordinate systems with different coordinate axes but with the same space and time origin, and the same sense of time flow, and which are mutually at rest; evidently all such transformations are possible. If we put $W=1$ and $\mathbf{v}=0$, we obtain transformations which connect two coordinate systems with different space and time origin, have the same sense of time flow, have coordinate axes in the same orientation, and are mutually at rest. Clearly all such have to be included in the relativity group. Finally, if we put $W=1$, $u_0=\mathbf{u}=0$, we obtain the transformations

$$x' = x + x_0\mathbf{v}, \qquad x_0' = x_0. \tag{10}$$

The formula (10) represents the transformation which connects two coordinate systems which have the same time origin and sense of time flow, which are moving uniformly with respect to each other with velocities $\mathbf{v}$ and $-\mathbf{v}$, and whose frames coincide for $x_0=x_0'=0$. These also have to be admitted. An easy calculation shows that these three classes of transformations generate the group of all transformations of the form (9). In other words, the physical assumptions we have made lead one to the group described by (9) as the relativity group. It is called the *orthochronous inhomogeneous Galilean group*, and the corresponding physical principles go under the name of *Galilean relativity*. If we assume $\det W = +1$ in (9), we obtain the usual inhomogeneous Galilean group. The transformation

$$\mathbf{x}' = \mathbf{x}, \qquad x_0' = -x_0, \tag{11}$$

represents *time inversion*. If we add this to the orthochronous group determined by (9), we generate a larger group called the *complete* inhomogeneous Galilean group.

When dealing with bodies and particles that move with very high velocities, these principles are violated quite decisively. The well known Michelson-Morley experiments led to the remarkable conclusion that in every inertial coordinate system, the velocity of propagation of electromagnetic and light signals in vacuum remained constant. It was Einstein who embodied this as a basic principle and deduced from it the consequence that the affine transformation (2) must be such that the same numerical value is obtained in both coordinate systems for the velocity of these signals. Choosing this numerical value to be the unit of time measurement, this means that for corresponding points of $\mathbf{R}^4$, the relation

$$(y_0-x_0)^2 - \sum_{j=1}^{3} (y_j-x_j)^2 = (y_0'-x_0')^2 - \sum_{j=1}^{3} (y_j'-x_j')^2 \tag{12}$$

must be satisfied. If we now substitute (2) in (12), we find, for the 4×4 matrix,

$$A = (a_{ij})_{i,j=0}^{3}, \tag{13}$$

the equation

$$A^t F A = F, \tag{14}$$

where F is the diagonal matrix:

$$F = \mathrm{diag}\{1,-1,-1,-1\}; \tag{15}$$

equivalently, the *linear* transformation induced in $\mathbf{R}^4$ by A preserves the *Minkowskian form*

$$\mathscr{G} : {x_0}^2-{x_1}^2-{x_2}^2-{x_3}^2. \tag{16}$$

The set of all A satisfying (14) is called the *Lorentz* group, and the set of all affine transformations (2), with A Lorentz is called the *complete inhomogeneous Lorentz* group. It can be shown (cf. Section 2) that $|a_{00}| \geq 1$ and that the set of all transformations with $a_{00} \geq 1$ is a subgroup, the *orthochronous* subgroup. Exactly as we showed in the Galilean case, it is possible to argue that the entire orthochronous Lorentz group must be admitted as the group of special relativity. To prove this, it is again sufficient to exhibit a generating class of transformations which actually arise. The transformations

$$\begin{aligned} \mathbf{x}' &= W\mathbf{x}+\mathbf{u}, \\ {x_0}' &= x_0+u_0, \end{aligned} \tag{17}$$

all arise explicitly; they connect inertial coordinate systems which are mutually at rest and which differ in their space and time origins and the orientations of their spatial coordinate axis but have the same sense of time flow. Consider, on the other hand, the transformations

$$\begin{aligned} {x_0}' &= (1-v^2)^{-1/2}x_0+v(1-v^2)^{-1/2}x_1, \\ {x_1}' &= v(1-v^2)^{-1/2}x_0+(1-v^2)^{-1/2}x_1, \\ {x_2}' &= x_2, \\ {x_3}' &= x_3, \end{aligned} \tag{18}$$

for values v with $-1<v<1$. Equation (18) represents, for fixed v, a homogeneous Lorentz transformation. The space-time point which corresponds to the point $(\tau,0,0,0)$ in the O'-system is represented by $(\tau(1-v^2)^{-1/2}, -\tau v(1-v^2)^{-1/2}, 0, 0)$ in the O-system. This shows that (18) represents the transformation which connects O and O' when O is moving relative to O' with uniform velocity v along the x_1-axis. Transformations of the form (18) must therefore be included in the space-time group. By symmetry, the transformations in which x_2 and x_3 play the role of x_1 are also admissible. It can be shown that these three classes of transformations, together with those defined by (17) and the time reversal (11), generate the complete inhomogeneous Lorentz group (cf. Section 2). In other words, in the special relativity of Einstein and Lorentz, the space-time group in question is the complete inhomogeneous Lorentz group.

It is usual in physical literature to do all this, not in natural units, but in the usual, say, C.G.S. units. If c denotes the velocity of light in vacuum as computed by inertial observers, then $x_0=ct$, where t is the time coordinate, and (18) will appear in the form

$$\begin{aligned} t' &= \frac{1}{\sqrt{1-v^2/c^2}}\,t+\frac{v/c^2}{\sqrt{1-v^2/c^2}}\,x_1, \\ x_1' &= \frac{v}{\sqrt{1-v^2/c^2}}\,t+\frac{1}{\sqrt{1-v^2/c^2}}\,x_1, \\ x_2' &= x_2, \\ x_3' &= x_3. \end{aligned} \tag{19}$$

When $v/c \to 0$, this formula becomes

$$\begin{aligned} t' &= t, \\ x_1' &= x_1+tv, \\ x_2' &= x_2, \\ x_3' &= x_3, \end{aligned} \tag{20}$$

which are the formulas expressing the same relation between the two frames in Galilean relativity. In other words, we may view the application of Galilean relativity as satisfactory when the velocities of the particles and bodies involved are small compared to c. For high-speed particles, the formulas (19) must be used, i.e., only Lorentz relativity is to be postulated.

These remarks conclude our extremely brief discussion of special relativity. The reader who is interested in a more detailed analysis of space and time is referred to Weyl's treatise [6].

The main problems of relativistic particle physics concern the various high-speed particles and their interactions. A general theory of such phenomena must combine the principles of quantum theory with those of special relativity to obtain the relevant equations of motion and to draw the necessary physical conclusions. In such generality these problems are still unsolved. In this chapter we shall be concerned only with the *free* particles and their equations. Before entering into a discussion we shall formulate mathematically the requirement that the physical description be relativistically invariant, i.e., the physical content of the systems under consideration be the same for all inertial coordinate systems.

Consider now a system $\mathfrak{S}$ which we want to describe in an invariant manner. We denote by G the space-time relativity group, namely, the group of all affine transformations $T_{O,O'}$ (2) which connect pairs of inertial systems. Each observer O associates with $\mathfrak{S}$ a Hilbert space $\mathscr{H}_O$ whose rays are identified by him with the *dynamical* or *Heisenberg* states (cf. Section 1 of Chapter VIII) of $\mathfrak{S}$. In the coordinate system O, $\mathfrak{S}$ is described by the determination of the set of all its dynamical states as the rays of

$\mathscr{H}_O$. In general, $\mathscr{H}_O$ will vary essentially with O. We shall formulate the relativistic invariance of (the description of) $\mathfrak{S}$ by the requirement that $\mathscr{H}_O$ be *independent* of O, i.e., the set of motions of $\mathfrak{S}$ is the same in all inertial coordinate systems. We shall write $\mathscr{H}$ for this Hilbert space. Even though $\mathscr{H}_O$ is the same in all inertial systems, the same physical state of $\mathfrak{S}$ may be described by two entirely different rays of $\mathscr{H}$ by two different observers, O, O'. Hence we assume the existence of a one-one mapping $t_{O,O'}$ of the set of all rays of $\mathscr{H}$ onto itself,

$$t_{O,O'} : r \rightarrow r', \tag{21}$$

such that r and $t_{O,O'}[r]$ represent (in O and O') the same state of $\mathfrak{S}$. Now, the transition probability between the dynamical states is an invariant numerical quantity, and hence $t_{O,O'}$ must preserve all the transition probabilities. This means that $t_{O,O'}$ is a physical symmetry in the sense of Chapter IV, and consequently is induced by a unitary or anti-unitary automorphism of $\mathscr{H}$ (Wigner's theorem). We write $D_{O,O'}$ for this symmetry. Clearly, we must have $D_{O,O}=1$ and $D_{O,O''}=D_{O'O''}D_{O,O'}$. If we write $T=T_{O,O'}$ for the element of G which connects O and O', D must depend only on T, as otherwise there would be an essential influence of $\mathfrak{S}$ on the coordinate system O. In this manner we obtain a representation $T \rightarrow D_T$ of the group G into the group of symmetries of $\mathscr{H}$ such that $t_{O,O'}$ is induced by $D_{T_{O,O'}}$ for all pairs O, O' of inertial coordinate systems. If G is a Lie group and G^0 is the component of identity of G, then we may assume that $T \rightarrow D_T$ is a homomorphism of G^0 into the projective group of $\mathscr{H}$ (cf. Wigner [1], [4] for a detailed discussion).

We thus see that associated with every relativistically invariant system $\mathfrak{S}$ there is a pair $(\mathscr{H},U)$ where $\mathscr{H}$ is the Hilbert space whose rays represent the dynamical states of $\mathfrak{S}$, and U is a (Borel) homomorphism of the connected component of the relativity group into the projective group of $\mathscr{H}$. We can replace U by a projective representation; any projective representation which induces the same homomorphism in the projective group may be used to study the behavior of the states under the relativity transformations. The physical interpretation of the representation will be along the same lines as in our analysis of systems covariant under configuration symmetries. We shall, of course, apply this only for the free particles, for which the interpretation is quite convincing.

It is easy to see how one may obtain the temporal evolution of $\mathfrak{S}$ from the representation U. To see this let us fix an inertial coordinate system O. We may then treat the rays of the Hilbert space $\mathscr{H}$ as describing the states of $\mathfrak{S}$, the state at time $t=0$ being identified with the same ray which describes the dynamical state. Consider now, for any real number a, the observer O' whose coordinate system is connected to that of O by the transformation T_{-a},

$$T_{-a} : x \rightarrow x',$$

where

$$x_0' = x_0 - a, \quad x_j' = x_j \qquad (1 \leq j \leq 3). \tag{22}$$

Then the map

$$a \to U_{T_{-a}} \qquad (-\infty < a < \infty) \tag{23}$$

gives the dynamical group of the system $\mathfrak{S}$.

We have already seen that the group $\mathbf{R}^1$ has only exact multipliers and so we can replace (23) by a representation $a \to V_a = \exp(-iaH)$. In view of our analysis in Chapter VIII it is natural to assume that H represents the *total energy* of the system. If U itself is a representation, then,

$$U_{T_a} = \exp(iaH). \tag{24}$$

Equation (24) can be generalized to describe the total linear momenta. Let $a \in \mathbf{R}^4$, let $a = (a_0, \mathbf{a})$, and let T_a be the transformation $x \to x'$, where

$$\begin{aligned} x_0' &= x_0 - a_0, \\ \mathbf{x}' &= \mathbf{x} - \mathbf{a}. \end{aligned}$$

Then there is a unitary representation $s \to W_s$ of $\mathbf{R}^1$ such that $s \to W_s$ induces the same homomorphism into the projective group as $s \to U_{T_{sa}}$ where $a = (0, \mathbf{a})$. We write $W_s = \exp(isB_{\mathbf{a}})$ and call $B_{\mathbf{a}}$ the total *linear momentum in the direction* $\mathbf{a}$. We may, of course, consider all the space-time translations at the same time by considering the restriction of U to the group of all space-time translations T_a. If we assume for the moment that the multipliers of the relativity group have exact restrictions to the translation subgroup, we can obtain a representation $a \to W_a$ ($a = (a_0, \mathbf{a})$) which can replace the restriction of U to the translation subgroup. By the general theory of representations of abelian groups, there exists a unique projection valued measure Q on the dual $\mathbf{P}^4$ of $\mathbf{R}^4$ such that

$$W_a = \int_{\mathbf{P}^4} \exp i\langle a, p \rangle dQ(p) \tag{25}$$

($\langle \cdot, \cdot \rangle$ is the duality between $\mathbf{R}^4$ and $\mathbf{P}^4$). The spectrum of Q is called the *spectrum of the system*. Notice that the energy and the momentum operators commute with one another. This leads, in the usual fashion (cf. Chapter VIII) to the law of conservation of the linear momenta.

It is clear that those forms of the representation U in which Q appears in its canonical form will be especially simple to analyze mathematically. Such descriptions are called *momentum space representations*.

Experiments involving the violation of parity have shown that one does not always have invariance of the physical content of theories with respect to the Lorentz transformations which do not belong to the connected component of the identity of G. Also it is well to remember that only elements of the connected component of the identity of G are represented by unitary operators. For the other elements it is necessary to use both unitary and anti-unitary operators. We shall examine this question later on.

We shall call a system *elementary* if the associated projective representation of the connected component of the relativity group is irreducible. In intuitive physical terms we can describe this by saying that starting from any state we may obtain all the states by Lorentz transformations and the principle of superposition. Two elementary systems may be said to be *physically equivalent* if there exists a unitary or anti-unitary isomorphism of the underlying Hilbert spaces which intertwines the associated projective representations. An *elementary particle in its free state* or a *free particle* is defined to be a physical equivalence class of elementary systems. One of the main results of this chapter is the classification of the elementary particles in terms of their mass and spin.

In classical and also in nonrelativistic quantum mechanics, the free particles do not play a very important role. However, the knowledge of the structure of the transformation laws for the states of free particles is very basic in most problems of particle theory. The reason for this is, roughly speaking, that most of the experiments involve scattering of various kinds of particles, and the incoming and outgoing are clearly those of free particles. It is because of this that a study of the transformation laws of free particles is quite fundamental.

2. THE LORENTZ GROUP

In this section we collect a number of facts involving the Lorentz group. The notations that will be used throughout this chapter are also introduced in this section.

We begin with the homogeneous Lorentz group, denoted by $\tilde{H}$. It is the group of all 4×4 matrices

$$L = (a_{ij})_{i,j=0}^{3} \tag{26}$$

such that

$$L^{t}FL = F, \tag{27}$$

where t denotes matrix transposition, and F is given by

$$F = \operatorname{diag}\{1,-1,-1,-1\}. \tag{28}$$

We write e_0, e_1, e_2, e_3 for the usual basis vectors so that $(x_0, x_1, x_2, x_3) = x_0e_0 + x_1e_1 + x_2e_2 + x_3e_3$. Each L induces a linear transformation of $\mathbf{R}^4$ and (27) is the condition that it preserve the quadratic form:

$$\mathscr{G} : {x_0}^2 - {x_1}^2 - {x_2}^2 - {x_3}^2. \tag{29}$$

The group $\tilde{H}$ is a closed subgroup of the group of all invertible linear transformations of $\mathbf{R}^4$, and as such is a lcsc group, even a Lie group. We shall examine its topology. Consider an $L \in \tilde{H}$. In view of (27),

$$\det L = \pm 1. \tag{30}$$

Further, L is an orthogonal matrix if and only if L preserves both the forms $x_1{}^2 + x_2{}^2 + x_3{}^2$ and $x_0{}^2$. Hence such an L is of the form

$$L = \begin{pmatrix} \pm 1 & 0 & 0 & 0 \\ 0 & & & \\ 0 & & W & \\ 0 & & & \end{pmatrix}, \tag{31}$$

where W is a 3×3 orthogonal matrix. We consider next a positive definite symmetric matrix L. Let $L^{1/2}$ denote the unique positive definite symmetric matrix whose square is L. We contend that if L is in $\tilde{H}$, so is $L^{1/2}$. In fact, from the equation $LFL = F$ we obtain $(L^{1/2})^2 = (FL^{-1/2}F)^2$, which shows, in view of the uniqueness of $L^{1/2}$, that $L^{1/2} = F(L^{1/2})^{-1}F$ or that $L^{1/2} \in \tilde{H}$. By a trivial induction we conclude that $L^{(m/2^n)} \in \tilde{H}$ for all integers m and n. From this observation we can determine the structure of the set of all positive definite Lorentz matrices. Consider, to this end, the mapping

$$S \to e^S$$

of the vector space of 4×4 real symmetric matrices into the space of symmetric positive definite matrices. It is well known (cf. Chevalley [1], pp. 14–15) that this is a homeomorphism. From what we saw just now, it can be asserted that whenever $e^S \in \tilde{H}$, $e^{(m/2^n)S} \in \tilde{H}$ for all m and n and hence that $e^{aS} \in \tilde{H}$ for all real a. If we differentiate the equation

$$(e^{aS})F(e^{aS}) = F$$

with respect to a, at $a = 0$, we obtain

$$SF + FS = 0 \qquad (S \text{ symmetric}). \tag{32}$$

Conversely, suppose that S is a 4×4 symmetric matrix satisfying (32). If we write $\varphi(a) = e^{aS}Fe^{aS}$, we find

$$\frac{d\varphi}{da}(a) = e^{aS}(SF + FS)e^{aS} = 0$$

for all a, proving that $\varphi(a)$ is a constant matrix; the equation $\varphi(0) = \varphi(1)$ tells us that $e^S \in \tilde{H}$. In other words, we have proved that for a 4×4 symmetric matrix S, $e^S \in \tilde{H}$ if and only if S satisfies (32). The set of all S

which satisfy (32) is a three-dimensional linear space $\mathfrak{p}$ with a basis given by

$$M_{01} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & & & \\ 0 & & 0 & \\ 0 & & & \end{pmatrix}, \qquad M_{02} = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & & & \\ 1 & & 0 & \\ 0 & & & \end{pmatrix},$$

$$M_{03} = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & & & \\ 0 & & 0 & \\ 1 & & & \end{pmatrix}. \tag{33}$$

If we write

$$H^p = \{L : L \in \tilde{H},\ L \text{ positive definite symmetric}\}, \tag{34}$$

then we may conclude that H^p is homeomorphic to the three-dimensional Euclidean space; $S \to e^S$ is a homeomorphism of $\mathfrak{p}$ onto H^p. In particular, $1 \in H^p$ and H^p is connected so that $H^p \subseteq H$, where H is the component of the identity of $\tilde{H}$. Consider now an arbitrary $L \in \tilde{H}$. Since $F = F^{-1}$, the equation $L^{-1}F(L^t)^{-1} = F$ shows that $(L^{-1})^t \in H$ and hence L^t itself belongs to H, i.e.,

$$\tilde{H}^t = \tilde{H}. \tag{35}$$

We now examine the *polar decomposition* of L, i.e., $L = L_1(L^tL)^{1/2}$, where L_1 is orthogonal. From (35) and our earlier discussion we know that $(L^tL)^{1/2} \in H^p$ and hence $L_1 \in \tilde{H}$. Hence L_1 must be of the form (31). In other words, every $L \in \tilde{H}$ can be uniquely written as

$$L = \varepsilon \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & & & \\ 0 & & \varepsilon' W & \\ 0 & & & \end{pmatrix} e^S, \tag{36}$$

where $W \in SO(3)$, $\varepsilon = \pm 1$, and $\varepsilon' = \pm 1$, and $S \in \mathfrak{p}$, i.e., S is symmetric and satisfies (32). Since the 0-0th entry of e^S is seen to be ≥ 1 by direct computation, it follows from (36) that for L,

$$\begin{aligned} |a_{00}| &\geq 1, \\ \varepsilon a_{00} &> 0. \end{aligned} \tag{37}$$

Moreover, $\det L = \varepsilon'^3$ so that

$$\varepsilon'(\det L) = 1. \tag{38}$$

Thus $\varepsilon=1$ and $\varepsilon'=1$ if and only if $a_{00}\geq 1$ and $\det L=+1$. Collecting all these facts, we have

Theorem 9.1. *The component H of the identity of $\tilde{H}$ is given by*

$$H = \{L : a_{00} > 0, \quad \det L = +1\}. \tag{39}$$

The map

$$W,S \rightarrow \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & & & \\ 0 & & W & \\ 0 & & & \end{pmatrix} e^S$$

is a homeomorphism of $SO(3)\times\mathfrak{p}$ onto H. For any Lorentz L, $|a_{00}|\geq 1$ and $\det L=\pm 1$. *$\tilde{H}$ has, besides H, exactly three connected components and these are the H-cosets of $\tilde{H}$ determined by the three elements I_s, I_t, and I_{st}, where I_s, I_t, and I_{st} are obtained from* (36) *by putting $S=0$, $W=1$, and for ε and ε' the respective pairs of values* $(1,-1)$, $(-1,-1)$, *and* $(-1,1)$.

We write K for the subgroup of H corresponding to $SO(3)$.

The Lie Algebra. Since $\tilde{H}$ is a linear group, we shall canonically identify its Lie algebra $\mathfrak{h}$ with a Lie subalgebra of matrices or endomorphisms of R^4, the bracket being the usual commutator $A, B\rightarrow[A,B]=AB-BA$. The exponential map then becomes the usual one, $S\rightarrow e^S$. A matrix $S\in\mathfrak{h}$ if and only if $e^{aS}\in\tilde{H}$ for all real a. Arguing as in (32), we find that

$$\mathfrak{h} = \{S : S^tF+FS = 0\}. \tag{40}$$

$\mathfrak{h}$ is six-dimensional. If we write $\mathfrak{k}$ for the set of matrices

$$\mathfrak{k} = \{S : S\in\mathfrak{h}, \quad Se_1 = 0\}, \tag{41}$$

then the elements S of $\mathfrak{k}$ are precisely those of the form

$$S = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & & & \\ 0 & & S_1 & \\ 0 & & & \end{pmatrix}, \tag{42}$$

where S_1 is skew symmetric:

$$S_1{}^t = -S_1.$$

$\mathfrak{k}$ is the subalgebra of $\mathfrak{h}$ corresponding to $SO(3)$ imbedded in $\tilde{H}$. $\mathfrak{h}$ is the direct sum of $\mathfrak{k}$ and $\mathfrak{p}$: $\mathfrak{h}=\mathfrak{k}+\mathfrak{p}$, $\mathfrak{k}\cap\mathfrak{p}=0$.

A useful basis for $\mathfrak{h}$ may be constructed in the following manner. Let e_{ij} denote the 4×4 matrix whose only nonzero entry is the i-jth, which is 1 ($i,j=0, 1, 2, 3$). Let

$$M_{ij} = -M_{ji} = e_{ij}-e_{ji} \qquad (i,j = 1, 2, 3, i \neq j), \tag{43}$$

$$M_{0i} = M_{0i} = e_{0i}+e_{i0} \qquad (i = 1, 2, 3). \tag{44}$$

Then M_{12}, M_{23}, M_{31}, M_{01}, M_{02}, and M_{03} span $\mathfrak{h}$. The commutation rules are given by:

$$\begin{aligned} [M_{ij},M_{jk}] &= M_{ik} && (i \neq j \neq k \neq i, \quad i, j, k = 1, 2, 3), \\ [M_{0i},M_{0j}] &= M_{ij} && (i \neq j, \quad i, j = 1, 2, 3), \\ [M_{0i},M_{jk}] &= 0 && (i \neq j \neq k \neq i, \quad i, j, k = 1, 2, 3), \\ [M_{ij},M_{0i}] &= -M_{0j} && (i \neq j, \quad i, j = 1, 2, 3). \end{aligned} \tag{45}$$

The Covering Group. We shall introduce in an explicit fashion the universal covering group of H. We consider the real four-dimensional space of 2×2 Hermitian complex matrices and the linear isomorphism of $\mathbf{R}^4$ with it given by

$$A : (x_0,x_1,x_2,x_3) \to x_0\sigma_0+x_1\sigma_1+x_2\sigma_2+x_3\sigma_3, \tag{46}$$

where σ_0 is the unit matrix, and σ_1, σ_2, and σ_3 are as in (55) of Chapter VI. The trace of this matrix is $2x_0$ and its determinant is $x_0{}^2-x_1{}^2-x_2{}^2-x_3{}^2$. Consider now the group

$$H^* = SL(2,\mathbf{C}), \tag{47}$$

and for any $m \in H^*$ the linear transformation

$$\delta(m) : \xi \to m\xi m^* \tag{48}$$

of the space of 2×2 Hermitian matrices. In view of (46), $\delta(m)$ induces a linear transformation in $\mathbf{R}^4$. Let us also write $\delta(m)$ for this transformation. From (48) we see that $m \to \delta(m)$ is a homomorphism of H^* into the group of invertible linear transformations of $\mathbf{R}^4$. Since $\det m=1$, $\det \xi = \det m\xi m^*$, so that the transformation $\delta(m)$ must preserve the quadratic form $x_0{}^2-x_1{}^2-x_2{}^2-x_3{}^2$. Thus $\delta(m) \in \tilde{H}$ and $m \to \delta(m)$ is a continuous homomorphism of H^* into $\tilde{H}$, hence into H. From (48) it follows easily that

$$\text{kernel } (\delta) = \{\pm 1\}. \tag{49}$$

δ is thus a local isomorphism and therefore its differential $\dot{\delta}$ induces an isomorphism of the Lie algebra of H^* into that of H. But the Lie algebra of H^*, being canonically isomorphic to the Lie algebra of 2×2 complex matrices of trace zero, is, as a real vector space, six-dimensional. Consequently, $\dot{\delta}$ is surjective. From this we infer at once that δ maps H^* onto the connected Lorentz group H:

$$\delta[H^*] = H. \tag{50}$$

The topology of H^* may be analyzed in the same way as the topology of H. The technical details are similar. Using polar decomposition of elements of H^*, we conclude that H^* is homeomorphic to the Cartesian product of K^*, its unitary subgroup, and the space of positive definite complex Hermitian matrices of determinant 1. Since $S \leftrightarrows e^S$ sets up a homeomorphism of the latter space with the three-dimensional real vector space of 2×2 Hermitian matrices of trace 0, we find that H^* is homeomorphic to $K^* \times \mathbf{R}^3$. In particular, H^* is simply connected. δ is thus a covering homomorphism of H^* onto H. We note that for an $m \in H^*$, $\delta(m)$ is orthogonal if and only if the map $\xi \to m\xi m^*$ preserves $\mathrm{tr}(\xi)$ as well as $\det(\xi)$. This can happen if and only if m is unitary. Thus

$$\delta[K^*] = K, \qquad K^* = \delta^{-1}(K). \tag{51}$$

The Lie algebra representation $\dot{\delta}$ corresponding to δ is easily computed. Let $\mathfrak{h}^*$ be the Lie algebra of 2×2 complex matrices of trace 0,

$$\mathfrak{h}^* = \{X : \mathrm{tr}\, X = 0, \quad X \text{ complex}\}. \tag{52}$$

Then

$$\dot{\delta}(X)\xi = X\xi + \xi X^*. \tag{53}$$

We take the basis $\{i\sigma_1, i\sigma_2, i\sigma_3, \sigma_1, \sigma_2, \sigma_3\}$ for $\mathfrak{h}^*$. An easy calculation gives:

$$\begin{aligned} &\dot{\delta}(i\sigma_1) = 2M_{23}, \quad \dot{\delta}(i\sigma_2) = 2M_{31}, \quad \dot{\delta}(i\sigma_3) = 2M_{12}, \\ &\dot{\delta}(\sigma_1) = 2M_{01}, \quad \dot{\delta}(\sigma_2) = 2M_{02}, \quad \dot{\delta}(\sigma_3) = 2M_{03}. \end{aligned} \tag{54}$$

From (54) we see that $\dot{\delta}$ maps Hermitian elements of $\mathfrak{h}^*$ into symmetric elements of $\mathfrak{h}$. Hence δ maps the set of positive definite elements of H^* onto H^p. Further, for a unitary $k \in K^*$, $\delta(k) \in K$. Hence, for a general element m of H^* of the form $m = kp$, k unitary, p positive-definite Hermitian, we have $\delta(m^*) = \delta(p)\delta(k^{-1}) = \delta(p)\delta(k)^t = (\delta(k)\delta(p))^t$ so that

$$\delta(m^*) = \delta(m)^t \qquad (m \in H^*). \tag{55}$$

On the other hand, let ζ be the reflection in R^4 given by

$$\zeta : x_0, x_1, x_2, x_3 \to x_0, x_1, -x_2, x_3. \tag{56}$$

ζ corresponds under the map (46), to the conjugation $\xi \to \bar{\xi}$ of Hermitian matrices. Since $\overline{m}\xi\overline{m}^* = (m\bar{\xi}m^*)^-$, we have

$$\delta(\overline{m}) = \zeta\delta(m)\zeta. \tag{57}$$

Combining (55) and (57), we find

$$\delta(m^t) = \zeta\delta(m^t)\zeta. \tag{58}$$

Representations of $SL(2,\mathbf{C})$ in Finite Dimensional Vector Spaces. The group H^* has no finite dimensional unitary representations other than the trivial representation, but it has a number of nonunitary representations in finite dimensional vector spaces. We shall now describe these.

We write $W=\mathbf{C}^2$ and W^j for the subspace of symmetric tensors of rank $2j$ ($j=0, \frac{1}{2}, 1, \cdots$) in the tensor product $W \otimes W \otimes \cdots \otimes W$ ($2j$ factors). The representation $m \rightarrow m \otimes m \otimes \cdots \otimes m$ leaves W^j invariant. We write D^j for the representation of H^* defined by W^j. D^j is already irreducible when restricted to the unitary group K^*. If $\{e_1, e_2\}$ is the standard basis for W, the matrix entries of the representation of H^* in $W \otimes \cdots \otimes W$, relative to the basis $\{e_{r_1} \otimes \cdots \otimes e_{r_N}\}$ ($N=2j$), are

$$f(r_1, \cdots, r_N, s_1, \cdots, s_N; m) = a_{r_1 s_1} \cdots a_{r_N s_N}, \tag{59}$$

where $m \in H^*$ is the matrix $(a_{rs})^2_{r,s=1}$. The f's are complex analytic functions on H^*. D^j is thus a complex analytic representation of H^* whose restriction to K^* is equivalent to the irreducible representation of K^* associated with j.

For any j, we may consider the representation $\bar{D}^j$ conjugate to D^j, i.e., we take a basis for W^j and conjugate the matrix representation obtained. $\bar{D}^j$ is also an irreducible representation of H^*. Its restriction to K^* is once again equivalent to the irreducible representations of K^* corresponding to j. No $\bar{D}^k$ is equivalent to any D^j except when $k = j = 0$.

It is known that for arbitrary j, j' ($j, j' = 0, \frac{1}{2}, 1, \cdots$) the representations

$$D^{j,j'} : D^j \otimes \bar{D}^{j'} \tag{60}$$

are all irreducible and that these exhaust, up to equivalence, the irreducible finite dimensional representations of H^*. For a proof the reader may refer to Weyl ([2], pp. 267–268).

The Inhomogeneous Group. The Lorentz group H acts in $\mathbf{R}^4$ as a linear transformation group and hence one may form the associated inhomogeneous group,

$$\tilde{G} = \tilde{H} \times' \mathbf{R}^4 \tag{61}$$

where we use the notation $\times'$ to denote the formation of the semidirect product. For (L,x) and (L',x') belonging to $\tilde{G}$,

$$(L,x)(L',x') = (LL', Lx'+x).$$

The group H^* may also be considered as acting on $\mathbf{R}^4$ through the action

$$mx = \delta(m)x \qquad (m \in H^*, x \in \mathbf{R}^4). \tag{62}$$

We write G^* for the associated semidirect product

$$G^* = H^* \times' \mathbf{R}^4. \tag{63}$$

The subgroup of $\tilde{G}$ given by

$$G = H \times' \mathbf{R}^4 \tag{64}$$

is the component of the identity in $\tilde{G}$, and the map

$$\delta : m, x \to \delta(m), x \tag{65}$$

is the covering homomorphism of G^* onto G; G^* is obviously simply connected.

The group $\tilde{G}$ is generated by G and the "inversions" $(I_s,0)$, $(I_t,0)$, and $(I_{st},0)$, where I_s is *space inversion*, I_t is *time inversion*, and I_{st} is *space-time inversion*: writing $(x_0,\mathbf{x})$ for $x_0e_0+x_1e_1+x_2e_2+x_3e_3$

$$\begin{aligned} I_s &: (x_0,\mathbf{x}) \to (x_0,-\mathbf{x}), \\ I_t &: (x_0,\mathbf{x}) \to (-x_0,\mathbf{x}), \\ I_{st} &: (x_0,\mathbf{x}) \to (-x_0,-\mathbf{x}). \end{aligned} \tag{66}$$

The elements I_τ ($\tau=0, s, t, st$; I_0 is the identity) form an abelian group having four elements, the *group of inversions*, which is denoted by H_{inv}:

$$H_{\text{inv}} = \{I_0,I_s,I_t,I_{st}\}. \tag{67}$$

The group H_{inv} is a discrete subgroup of the homogeneous Lorentz group $\tilde{H}$, and $\tilde{H}$ is the semidirect product of H_{inv} and H:

$$\tilde{H} = H_{\text{inv}}H, \qquad H \cap H_{\text{inv}} = \{1\}; \tag{68}$$

the action of H_{inv} on H is of course through inner automorphisms.

For any τ, $L \to I_\tau L I_\tau^{-1}$ is an involutive automorphism of H, and the assignment which sends τ to this automorphism is a homomorphism. As H^* is the universal covering group of H, there exists for each τ, a unique automorphism $h \to h_\tau$ of H^* such that

$$\delta(h_\tau) = I_\tau \delta(h) I_\tau^{-1} \tag{69}$$

for all h. $h \to h_\tau$ is involutive and the assignment s, which sends τ to the automorphism $h \to h_\tau$, is necessarily a homomorphism. We may thus form the semidirect product

$$\tilde{H}^* = H_{\text{inv}} \times_s H^*. \tag{70}$$

We continue to write δ for the map

$$I_\tau, h \to \delta(h) I_\tau;$$

δ is a homomorphism of $H_{\text{inv}} \times_s H^*$ onto $\tilde{H}$.

The complete inhomogeneous group $\tilde{G}$ is the semidirect product of G and H_{inv}; each I_τ gives rise to the automorphism

$$(L,x) \to (I_\tau L I_\tau, I_\tau x)$$

of $\tilde{G}$. The corresponding map

$$(h,x) \to (h_\tau, I_\tau x)$$

is seen to be an automorphism of G^*, and the assignment s, which sends τ to this automorphism, is once again a homomorphism. We may thus form the semidirect product $\tilde{G}^*$,

$$\tilde{G}^* = H_{\text{inv}} \times_s G^*.$$

The group $\tilde{G}^*$ has a natural homomorphism δ onto $\tilde{G}$:

(71) $$\delta : I_\tau, h, x \to (\delta(h)I_\tau, x).$$

It is useful to determine explicitly the automorphisms $h \to h_\tau$. For $\tau = st$, $I_{st}LI_{st}^{-1} = L$ so that

(72) $$h_{st} = h.$$

Suppose now that $\tau = t$ so that I_τ is time inversion. It is clear that $I_t = -F$, F being the matrix of the fundamental form. Hence $\delta(h^{-1*}) = (\delta(h)^{-1})^t$ (cf. (55)) $= F\delta(h)F$ so that

(73) $$h_t = h^{-1*}.$$

As $I_sI_t = I_{st}$, we have

(74) $$h_s = h_t.$$

Orbits in Momentum Space. The space $\mathbf{R}^4$ is a locally compact abelian group under addition and hence possesses a dual group. We identify this group with the vector space dual of $\mathbf{R}^4$. To keep things explicitly separated we write $\mathbf{P}^4$ for the vector space dual of $\mathbf{R}^4$. Elements of $\mathbf{P}^4$ will be written

(75) $$p = (p_0, p_1, p_2, p_3) = (p_0, \mathbf{p}),$$

and we shall define the canonical duality between $\mathbf{R}^4$ and $\mathbf{P}^4$ by

(76) $$\{x,p\} = x_0p_0 - x_1p_1 - x_2p_2 - x_3p_3.$$

We associate with each $p \in \mathbf{P}^4$ the character

$$\hat{p} : x \to \exp i\{x,p\}$$

of $\mathbf{R}^4$. The mapping $p \to \hat{p}$ maps $\mathbf{P}^4$ isomorphically onto the character group of $\mathbf{R}^4$. We thus identify $\mathbf{P}^4$ with the character group of $\mathbf{R}^4$.

Since each element L of $\tilde{H}$ is a matrix, it acts in $\mathbf{R}^4$ as well as in $\mathbf{P}^4$. In view of the definition of $\{.,.\}$ and $\hat{p}$, we have:

(77) $$\{Lx, Lp\} = \{x,p\},$$

(78) $$\hat{p}(L^{-1}x) = (Lp)\hat{}\,(x).$$

We now compare (77) and (78) with (127) of Chapter VI, and conclude that $p \rightarrow Lp$ gives the adjoint action of L on $\mathbf{P}^4$. It is customary to refer to $\mathbf{P}^4$ as *momentum space.*

The group $\tilde{H}$ acts on $\mathbf{P}^4$ and converts $\mathbf{P}^4$ into a $\tilde{H}$-space. $\mathbf{P}^4$ then breaks up into $\tilde{H}$-orbits and it is necessary to compute these. It is simplest to deal with the H-orbits. We introduce the following subsets of $\mathbf{P}^4$:

$$\begin{aligned} X_m^+ &= \{p : p_0^2-p_1^2-p_2^2-p_3^2 = m^2, p_0 > 0\}, \\ X_m^- &= \{p : p_0^2-p_1^2-p_2^2-p_3^2 = m^2, p_0 < 0\}, \end{aligned} \tag{79}$$

where m is real and ≥ 0. We write

$$X_{00} = \{0\}.$$

Also, for any $m>0$, we write

$$Y_m = \{p : p_0^2-p_1^2-p_2^2-p_3^2 = -m^2\}.$$

$X_m^+ \cup X_m^-$ and Y_m are obviously invariant under the whole homogeneous Lorentz group $\tilde{H}$, in particular, under H. We claim that X_m^+ and X_m^- $(m \geq 0)$ are invariant under H. If $m \geq 0$, the consideration of the mapping $(p_1,p_2,p_3) \rightarrow \pm(p_1^2+p_2^2+p_3^2)^{1/2}$ shows that X_m^+ and X_m^- are the connected components of $X_m^+ \cup X_m^-$. Since H is connected, it must leave each of them invariant. For $m>0$, X_m^+ and X_m^- are homeomorphic to $\mathbf{R}^3$; for $m=0$, to $\mathbf{R}^3-\{(0,0,0)\}$. Thus $X_m^\pm$ $(m \geq 0)$, Y_m $(m>0)$, and X_{00} are all H-invariant, and $\mathbf{P}^4$ is the union of these.

We shall next prove that H acts *transitively* on each of these. For real x, y let

$$L_{x,y} = \begin{pmatrix} y & x & 0 & 0 \\ x & y & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \tag{80}$$

If $y>0$ and $y^2-x^2=1$, it is clear that $L_{x,y}$ belongs to H. Consider now X_m^+ with $m>0$ and a point p on it. The point $(m,0,0,0)$ lies on X_m^+. Choose now $x=(p_0^2-m^2)^{1/2}/m$, $y=p_0/m$ in (80); since $p_0>0$ and $p_0^2-m^2=p_1^2+p_2^2+p_3^2$, x and y are real and $y>0$. Then $L=L_{x,y}$ sends $(m,0,0,0)$ to a point of the form $(p_0,p_1',0,0)$. Now $p_1'^2=p_1^2+p_2^2+p_3^2$ so that we can find a rotation $R \in K$ sending $(p_0,p_1',0,0)$ to (p_0,p_1,p_2,p_3). This shows that RL sends $(m,0,0,0)$ to p. H is thus transitive on X_m^+. For X_m^- we consider $(-m,0,0,0)$ and use a similar argument. Consider next X_0^+ and a point $p \in X_0^+$. $(1,1,0,0) \in X_0^+$ and, if we choose $x=(p_0^2-1)/2p_0$ and $y=(p_0^2+1)/2p_0$ in (80), $L=L_{x,y}$ sends $(1,1,0,0)$ into $(p_0,p_0,0,0)$. Once again, there is a rotation $R \in K$ such that RL sends $(1,1,0,0)$ to p. A similar argument, using $(-1,-1,0,0)$ instead of $(1,1,0,0)$

establishes the transitivity of H on X_0^-. The transitivity of H on X_{00} is trivial.

We now take up the Y_m $(m>0)$. Let $p=(p_0,p_1,p_2,p_3)\in Y_m$; p_0 need not be >0. If we choose $x=p_0/m$ and $y=(p_0^2+m^2)^{1/2}/m$, then $L=L_{x,y}\in H$ and sends the point $(0,m,0,0)$ of Y_m into $(p_0,my,0,0)$; as $m^2y^2=p_1^2+p_2^2+p_3^2$, there is a rotation $R\in K$ such that RL sends $(0,m,0,0)$ to p. This proves that Y_m is transitively acted on by H.

The stability subgroup of H at $(m,0,0,0)$ or $(-m,0,0,0)$ $(m>0)$ is the rotation subgroup K of H. The stability subgroup at $(0,m,0,0)$ is the subgroup of H consisting of all matrices L of the form

$$L=\begin{pmatrix} a_{00} & 0 & a_{02} & a_{03} \\ 0 & 1 & 0 & 0 \\ a_{20} & 0 & a_{22} & a_{23} \\ a_{30} & 0 & a_{32} & a_{33} \end{pmatrix},$$

where

$$a_{00}>0$$

and the 3×3 matrix $(a_{ij})_{i,j=0,2,3}$ has determinant 1 and preserves the form

$$\mathscr{G}' : x_0^2-x_2^2-x_3^2.$$

In our usual notation (cf. Chapter V), we denote these by $H_{(\pm m,0,0,0)}$ and $H_{(0,m,0,0)}$. The first two are compact and connected. The last is not compact. All are semisimple. For the action of the covering group H^* on $\mathbf{P}^4$, the orbits are the same. For the stability groups, we have:

$$H^*_{(\pm m,0,0,0)}=K^*. \tag{81}$$

For $X_0^{\pm}$ the nature of the stability subgroups changes radically. It is most convenient to construct these corresponding to the two points $(\pm 1,0,0,\pm 1)$. We consider X_0^+ and the pre-image in H^* of the stability group at $(1,0,0,1)$:

$$H^*_{(1,0,0,1)}=\delta^{-1}(H_{(1,0,0,1)}). \tag{82}$$

Using (46) we find that $m\in H^*_{(1,0,0,1)}$ if and only if

$$m\begin{pmatrix}1 & 0\\ 0 & 0\end{pmatrix}m^*=\begin{pmatrix}1 & 0\\ 0 & 0\end{pmatrix}. \tag{83}$$

$H^*_{(1,0,0,1)}$ is thus the group E^* of all matrices m of the form

$$m_{z,a}=\begin{pmatrix}z & a\\ 0 & z^{-1}\end{pmatrix} \qquad (a,z\in\mathbf{C},\ |z|=1). \tag{84}$$

$\{m_{1,a} : a \in \mathbf{C}\}$ is a normal subgroup isomorphic to the plane, $\{m_{z,0} : |z|=1\}$ is isomorphic to T, and E^* is the semidirect product of these two where the action of z on a is given by $a \to z^2 a$. Therefore $E^*/\{\pm 1\}$ (1 in the denominator is the 2×2 unit matrix) is isomorphic canonically to the *Euclidean group in the plane.* In other words $H_{(1,0,0,1)}$ is isomorphic to the group of Euclidean motions of the plane R^2. The groups $H^*_{(1,0,0,1)}$ and $H_{(1,0,0,1)}$ are not semisimple. They are, in fact, solvable. The calculations are similar for the case of X_0^-; the stability subgroups are the same.

We note that our argument showing that H is transitive on X_m^+ established something more. It actually showed that given any $p \in X_m^+$ there is a rotation $R \in K$ and a Lorentz transformation $L_{x,y}$ $(x,y>0$, cf. (80)), such that $RL_{x,y}$ sends $(m,0,0,0)$ to p. As the stability subgroup at $(m,0,0,0)$ is K, we conclude from this that any $L \in H$ can be written as

$$L = RZS, \tag{85}$$

where R and S are rotations, and $Z = L_{x,y}$ for suitable $x \geq 0$, $y > 0$ with $y^2 - x^2 = 1$. If we define

$$v = x/y, \tag{86}$$

then $L_{x,y}$ becomes the transformation connecting two inertial frames, the second of which is moving uniformly with respect to the first with velocity v (cf. (18)). This shows that the general Lorentz transformation is a product of the rotations and transformations to moving frames of the form (18). We have already used this fact in Section 1.

The Invariant Measures on the Orbits. All the H-orbits admit invariant Borel measures. This is obvious for $X_m^{\pm}$ and the Y_m $(m>0)$ because both the group H and the stability groups are semisimple and hence unimodular. For $X_0^{\pm}$ we note that the Euclidean group is unimodular so that each of $X_0^{\pm}$ also admits an invariant measure. We shall now calculate these measures explicitly.

Choose a real number $a>0$. For any $\xi > a$ let D_ξ be the H-invariant open set $D_\xi = \{p : p_0 > 0, a < p_0^2 - p_1^2 - p_2^2 - p_3^2 < \xi\}$. Let

$$F(f,\xi) = \int_{D_\xi} f(p)dp \qquad (f \in C_c(P^4)).$$

Since det $L=1$ for $L \in H$, $F(f^L,\xi) = F(f,\xi)$ for all $\xi > a$ and $L \in H$. Moreover, if $f \geq 0$, $F(f,\xi)$ increases with ξ. Hence, on writing

$$J_\xi(f) = \frac{d}{d\xi} F(f,\xi) \qquad (\xi > a),$$

we see that $f \to J_\xi(f)$ is an H-invariant nonnegative linear functional on $C_c(\mathbf{P}^4)$. It therefore determines an H-invariant Borel measure on $\mathbf{P}^4$. On the other hand, the transformation

$$S : (p_0,p_1,p_2,p_3) \to (p_0^2 - p_1^2 - p_2^2 - p_3^2, p_1, p_2, p_3)$$

is one-one on D_ξ so that we also have

$$F(f,\xi) = \int_a^\xi \left(\int_{\mathbf{P}^3} \frac{f((u+|\mathbf{p}|^2)^{1/2}, \mathbf{p})}{2(u+|\mathbf{p}|^2)^{1/2}} \, d\mathbf{p} \right) du$$

($|\mathbf{p}|^2 = p_1^2 + p_2^2 + p_3^2$). From this we obtain

$$J_\xi(f) = \int_{\mathbf{P}^3} \frac{f((\xi+|\mathbf{p}|^2)^{1/2}, \mathbf{p})}{2(\xi+|\mathbf{p}|^2)^{1/2}} \, d\mathbf{p}.$$

This formula shows that the measure defined by J_{m^2} is actually supported by the set $X_m{}^+$ and is thus the invariant measure on the orbit $X_m{}^+$. We write $\alpha_m{}^+$ for this measure. Thus

$$\int_{X_m^+} f d\alpha_m{}^+ = \int_{\mathbf{P}^3} \frac{f((m^2+|\mathbf{p}|^2)^{1/2}, \mathbf{p})}{2(m^2+|\mathbf{p}|^2)^{1/2}} \, d\mathbf{p}. \tag{87}$$

Note that as $m \to 0+$, the integral on the right remains finite, since the function $\mathbf{p} \to |\mathbf{p}|^{-1}$ is *integrable around the origin* in the space $\mathbf{P}^3$. Hence the same formula defines the invariant measure on $X_0{}^+$ also. Note that $X_0{}^+$ is not closed in $\mathbf{P}^4$ and yet the invariant measure on $X_0{}^+$ is actually a Borel measure on $\mathbf{P}^4$.

A similar computation leads to the following formula for the invariant measure $\alpha_m{}^-$ on $X_m{}^-$:

$$\int_{X_m^-} f d\alpha_m{}^- = \int_{\mathbf{P}^3} \frac{f(-(m^2+|\mathbf{p}|^2)^{1/2}, \mathbf{p})}{2(m^2+|\mathbf{p}|^2)^{1/2}} \, d\mathbf{p}. \tag{88}$$

We may follow an analogous method for determining the invariant measures β_m on the orbits Y_m. We choose a and ξ such that $a < \xi < 0$ and define D_ξ' as the open set $\{p : a < p_0{}^2 - p_1{}^2 - p_2{}^2 - p_3{}^2 < \xi\}$. The transformation S is not one-one on D_ξ' but one-one separately on $D_\xi' \cap \{p : p_0 > 0\}$ and $D_\xi' \cap \{p : p_0 < 0\}$. Hence:

$$\int_{D_\xi'} f dp = \int_a^\xi \left(\int_{|\mathbf{p}|^2 > -u} \frac{f((u+|\mathbf{p}|^2)^{1/2}, \mathbf{p}) + f(-(u+|\mathbf{p}|^2)^{1/2}, \mathbf{p})}{2(u+|\mathbf{p}|^2)^{1/2}} \, d\mathbf{p} \right) du.$$

We may now proceed as in the earlier case and obtain the formula

$$\int_{Y_m} f d\beta_m = \int_{|\mathbf{p}| > m} \frac{f((-m^2+|\mathbf{p}|^2)^{1/2}, \mathbf{p}) + f(-(-m^2+|\mathbf{p}|^2)^{1/2}, \mathbf{p})}{2(-m^2+|\mathbf{p}|^2)^{1/2}} \, d\mathbf{p} \tag{89}$$

for all $f \in C_c(\mathbf{P}^4)$.

The foregoing observations are collected together in the following theorem.

Theorem 9.2. *The group H acts transitively on $X_m{}^+$, $X_m{}^-$ ($m \geqslant 0$), X_{00}, and Y_m ($m > 0$) which, as m varies, exhaust the orbits of H in momentum space. The stability subgroup of H^* at ($\pm m$,0,0,0) is the unitary subgroup*

K^*, while the stability subgroup at $(\pm 1,0,0,\pm 1)$ is the subgroup E^* (85) *of H^*. Each orbit admits invariant measures. If we use p_1, p_2, p_3 as global coordinates on $X_m{}^\pm$ and Y_m, the invariant measures $\alpha_m{}^\pm$ and β_m are given by*

$$d\alpha_m{}^\pm \sim \frac{dp_1dp_2dp_3}{2(m^2+p_1{}^2+p_2{}^2+p_3{}^2)^{1/2}} \qquad (m \geq 0),$$

$$d\beta_m \sim \frac{dp_1dp_2dp_3}{2(-m^2+p_1{}^2+p_2{}^2+p_3{}^2)^{1/2}} \qquad (m > 0, \quad |\mathbf{p}| > m).$$

3. THE REPRESENTATIONS OF THE INHOMOGENEOUS LORENTZ GROUP

The inhomogeneous Lorentz group G as well as its covering group G^* are semidirect products. Moreover, from theorem 7.40 we conclude that all the multipliers of G^* are exact and also that every irreducible projective representation of G is induced canonically from a unique irreducible representation of G^*. The group G^* is a semidirect product and to it we apply the theory of Chapter VI, especially theorem 6.24. The orbits of H^* in $\mathbf{P}^4$ have been calculated in Section 2 and it is obvious that the orbit structure is smooth. Furthermore, the stability subgroups of H^* associated with these orbits are described in (81) and (84). Theorem 6.24 therefore enables us to describe the irreducible unitary representations of G^* completely. In what follows we shall carry out such a description, arranged according to the various orbits.

The Orbit X_{00}. Since $X_{00}=\{p : p=0\}$, the stability subgroup is the whole group H^*. The irreducible representations of G^* associated with X_{00} are obtained by selecting an irreducible unitary representation π of H^* and defining the representation $U^{00,\pi}$ of G^* by:

$$U^{00,\pi} : h, x \to \pi(h) \qquad (h \in H^*). \tag{90}$$

The Orbits $X_m{}^\pm$ $(m>0)$. We select $(\pm m,0,0,0)$ as representative points of $X_m{}^\pm$. The stability subgroup of H^* is the unitary subgroup K^* whose irreducible unitary representations are the representations D^j $(j=0,\frac{1}{2},1,\ldots)$ described in Chapter VI. Let us choose a strict $(H^*,X_m{}^\pm)$-cocycle φ^j whose values are unitary operators in the space W^j of D^j such that φ^j defines the representation D^j of K^* at $(\pm m,0,0,0)$. Let

$$\mathscr{H}^{m,\pm,j} = \mathscr{L}^2(X_m{}^\pm,W^j,\alpha_m{}^\pm). \tag{91}$$

For $x \in \mathbf{R}^4$ and $h \in H^*$ we define the operator $U^{m,\pm,j}_{h,x}$ of $\mathscr{H}^{m,\pm,j}$ by

$$(U^{m,\pm,j}_{h,x}f)(p) = \exp i\{x,p\}\varphi^j(h,\delta(h)^{-1}p)f(\delta(h)^{-1}p); \tag{92}$$

here $\{.,.\}$ is defined as in (76) and $\delta(h)$ is the Lorentz matrix corresponding to $h \in H^*$. In view of our discussion in Chapter VI we know that

$$U^{m,\pm,j} : h, x \rightarrow U^{m,\pm,j}_{h,x} \tag{93}$$

is the irreducible representation of G^* corresponding to the orbit $X_m{}^{\pm}$ and the representation D^j of K^*.

The Orbits $X_0{}^{\pm}$. The stability group at the points $(\pm 1,0,0,\pm 1)$ is the set of all matrices ($z, a \in \mathbf{C}$, $|z| = 1$)

$$m_{z,a} = \begin{pmatrix} z & a \\ 0 & z^{-1} \end{pmatrix}.$$

We have written E^* for this group. Our first objective is to determine the irreducible unitary representations of E^*. Now, E^* is itself a semidirect product so that this can be done with the help of the theory of Chapter VI. We write $A = \{m_{1,a}\}$, $T' = \{m_{z,0}\}$. Clearly

$$\begin{aligned} m_{z,a} &= m_{1,az}m_{z,0}, \\ m_{z,0}m_{1,a}m_{z,0}^{-1} &= m_{1,z^2a}. \end{aligned} \tag{94}$$

Thus, with respect to the action $z, a \rightarrow z^2a$ of T' in A, E^* becomes identified with the semidirect product of A and T'.

For each complex number b,

$$y_b : m_{1,a} \rightarrow \exp i \operatorname{Re}(ba^*)$$

is a character of A, and $b \rightarrow y_b$ is an isomorphism of the additive complex number group onto $\hat{A}$. The adjoint action (cf. (127) of Chapter VI) of T' on A is given by

$$m_{z,0} : y_b \rightarrow y_{z^2b}$$

and the T'-orbits in A are

$$C_0 = \{y_0\}$$

and

$$C_\rho = \{y_b : |b| = \rho\} \qquad (\rho > 0).$$

The stability subgroup at y_0 is T' itself; at any point of C_ρ it is $\{m_{\pm 1,0}\}$.

The representations of E^* associated with y_0 are simply those of T' lifted to E^*. Thus, for each integer n,

$$\pi_n : m_{z,a} \rightarrow z^n \qquad (n = 0, \pm 1, \pm 2, \cdots) \tag{95}$$

is a unitary irreducible representation of E^*. On the other hand, if $\rho > 0$, we note that the stability group associated with C_ρ has two one-dimensional irreducible representations—the trivial one, and the one which sends $m_{\pm 1,0}$ to ± 1. Corresponding to these, there are two inequivalent irreducible representations π_ρ and π_ρ' of E^*. Since π_ρ is associated with the trivial representation of the stability group corresponding to C_ρ, its description is immediate. For describing π_ρ' we must first determine a strict (T',C_ρ)-cocycle with values in T which defines the homomorphism $m_{\pm 1,0} \to \pm 1$ at a point of C_ρ. It is obvious that the cocycle

$$m_{z,0}, y \to z \qquad (y \in C_\rho) \tag{96}$$

has the required properties. We thus obtain the following description of π_ρ and π_ρ'. Both of them act in

$$\mathscr{K} = \mathscr{L}^2(T,dt), \tag{97}$$

and

$$(\pi_\rho(m_{z,a})f)(t) = \exp[i\rho \operatorname{Re}(tz^{-1}a^*)]f(z^{-2}t), \tag{98}$$

$$(\pi_\rho'(m_{z,a})f)(t) = \exp[i\rho \operatorname{Re}(tz^{-1}a^*)]zf(z^{-2}t); \tag{99}$$

we have simply taken into account the formulas (94) and applied (132) of Chapter VI.

Having determined the irreducible representations of E^* we can now obtain the representations of G^* which are associated with the orbits $X_0{}^{\pm}$. We select for $n = 0, \pm 1, \pm 2, \cdots$ a strict $(H^*,X_0{}^{\pm})$ cocycle with values in T, say φ^n, which defines the representation:

$$\pi_n : m_{z,a} \to z^n$$

at $(\pm 1,0,0,\pm 1)$. Then, in the Hilbert space

$$\mathscr{H}^{\pm,n} = \mathscr{L}^2(X_0{}^{\pm},\alpha_0{}^{\pm}), \tag{100}$$

we have the representation $U^{\pm,n}$ of G^* given by:

$$(U^{\pm;n}_{h,x}f)(p) = \exp i\{x,p\}\varphi^n(h,\delta(h)^{-1}p)f(\delta(h)^{-1}p). \tag{101}$$

Furthermore, choose for each ρ, strict $(H^*,X_0{}^{\pm})$-cocycles φ^ρ and $\varphi^{\rho,'}$ which take values in the unitary group of $\mathscr{K} = \mathscr{L}^2(T,dt)$, and which define, at $(\pm 1,0,0,\pm 1)$, the representations π_ρ and π_ρ' of E^*. The corresponding representations $U^{\pm,\rho}$, $U^{\pm,\rho,'}$ of G^* act in

$$\mathscr{H}^{\pm,\rho} = \mathscr{L}^2(X_0{}^{\pm},\mathscr{K},\alpha_0{}^{\pm}) \tag{102}$$

and

$$(U^{\pm;\rho}_{h,x}f)(p) = \exp i\{x,p\}\varphi^\rho(h,\delta(h)^{-1}p)f(\delta(h)^{-1}p), \tag{103}$$

$$(U^{\pm;\rho,'}_{h,x}f)(p) = \exp i\{x,p\}\varphi^{\rho,'}(h,\delta(h)^{-1}p)f(\delta(h)^{-1}p). \tag{104}$$

The Orbits Y_m ($m > 0$). The stability subgroup at $(0,m,0,0)$ is a semisimple Lie group but is not compact. Its unitary representations other than the trivial representation are all infinite dimensional. The general theory associates with any one of these, say π, a representation $U^{im,\pi}$ of G^*. We shall not bother to write these down since, as we shall see presently, they do not seem to be useful in describing physical particles.

We shall now take up the question of classifying the representations obtained into physical equivalence classes. It is clear that two representations of G^* are physically equivalent if and only if there exists a unitary or anti-unitary isomorphism which intertwines them. Now we have seen in the course of the proof of lemma 8.3 that, given an arbitrary representation π of a group, there is a canonical anti-unitary isomorphism of the Hilbert space of the representation π onto its dual which intertwines π with its contragredient. It follows, therefore, that two representations of G^* are physically equivalent if and only if they are either equivalent or mutually contragredient. The case of contragredience is settled by the following lemma.

Lemma 9.3. *Let $F \times' B$ be a semidirect product and let $0 \subseteq \hat{B}$ be an F-orbit under the adjoint action. Let F_y be the stability subgroup of F at $y \in O$. Then $F_{y_0^{-1}} = F_{y_0}$, and if O^{-1} is the orbit of y_0^{-1}, then for any representation π of F_{y_0}, the representations of $F \times' B$ associated with (O,π) and (O^{-1},π^0) are mutually contragredient, π^0 being the representation contragredient to π.*

Proof. Let π act in a Hilbert space $\mathscr{K}$. We select a finite measure α on O, quasi invariant under the F action, and define, for $h \in F$, $r_h = d\alpha/d\alpha^{(h^{-1})}$. Then, the representation U of $F \times' B$ which is associated with (O,π) can be written as

$$(U_{h,b}f)(y) = \{r_h(h^{-1} \cdot y)\}^{1/2} y(b)\varphi(h,h^{-1}y)f(h^{-1} \cdot y) \qquad (y \in O),$$

where φ is a strict (F,O)-cocycle with values in the unitary group of $\mathscr{K}$; U acts in $\mathscr{H} = \mathscr{L}^2(O,\mathscr{K},\alpha)$. Now, F acts as a group of automorphisms of the group $\hat{B}$. From this it follows that $F_{(y_0^{-1})} = F_{y_0}$ and that the map $y \to y^{-1}$ transfers the measure α on O to a quasi-invariant measure β on O^{-1}. Choose any anti-unitary automorphism S of $\mathscr{K}$ and define J by

$$(Jf)(y) = Sf(y^{-1}) \qquad (y \in O^{-1}).$$

J is clearly an anti-unitary isomorphism of $\mathscr{H}$ onto $\mathscr{H}^0 = \mathscr{L}^2(O^{-1},\mathscr{K},\beta)$. We compute $J \circ U \circ J^{-1} = V$. A brief calculation shows that, for $f^0 \in \mathscr{H}^0$. we have

$$(V_{h,b}f^0)(y) = \{r_h(h^{-1} \cdot y^{-1})\}^{1/2} y(b) S\varphi(h,h^{-1} \cdot y^{-1})S^{-1}f^0(h^{-1} \cdot y).$$

It is obvious that $z \to r_h(z^{-1})$ is a version of $d\beta/d\beta^{(h^{-1})}$ and that $h, z \to S\varphi(h,z^{-1})S^{-1}$ is a strict (F,O^{-1})-cocycle defining the representation $S \circ \pi \circ S^{-1}$ at y_0^{-1}. Hence V is the representation associated with O^{-1} and $S \circ \pi \circ S^{-1}$. Since $S \circ \pi \circ S^{-1} \simeq \pi^0$ and $J \circ U \circ J^{-1} \simeq U^0$ (the contragredient of U), the lemma is proved.

Applying this lemma we see that the physical equivalence classes corresponding to positive m are $\{U^{m,+,j}, U^{m,-,j}\}$. For the representations associated with $X_0^{\pm}$ it is clear in the first place that $U^{+,n}$ and $U^{-,-n}$ are physically equivalent. On the other hand, lemma 9.3 applied to E^* shows that the representations π_ρ and π_ρ' are both self-contragredient, so that we conclude that $U^{+,\rho}$ and $U^{-,\rho}$ as well as $U^{+,\rho,'}$ and $U^{-,\rho,'}$ are physically equivalent. We can now formulate the main result of this chapter. It was obtained by Wigner in his 1939 paper.

Theorem 9.4 (*Wigner* [1]). *Any irreducible representation of G^* induces a projective representation of G which is irreducible and, conversely, given an irreducible projective representation of G there is exactly one irreducible representation of G^* which induces it. An arbitrary irreducible representation of G^* is equivalent to one associated with some H^*-orbit in P^4 and some irreducible representation of the stability subgroup of H^* at some point of this orbit. The irreducible representations $U^{m,+,j}$ $(j=0, \frac{1}{2}, \cdots)$, $U^{+,\pm n}$ $(n=0, 1, 2, \cdots)$, $U^{+,\rho}$ and $U^{+,\rho,'}$ $(\rho>0)$ are physically inequivalent and an irreducible representation of G^* which is associated with one of the orbits $X_m^{\pm}$ $(m \geq 0)$ is physically equivalent to exactly one of these.*

Not all irreducible representations of G^* can be used to represent free particles. Consider first of all the representations $U^{00,\pi}$ associated with the trivial orbit X_{00}. In every one of these, the space-time translations are represented by the identity operator. This implies that the energy and momentum observables vanish identically. Consequently, all these representations have to be given up, with the exception of the trivial one-dimensional representation of G^*, in which *every* Lorentz transformation goes over to the identity. It represents the *vacuum*.

We shall next examine an irreducible representation U of G^* associated with one of the nontrivial orbits $X_m^{\pm}$ and Y_m. Let Q_0 be the energy and Q_1, Q_2, Q_3 the linear momenta along the three axes in a given coordinate system. Since the space-time translation by $x \in \mathbf{R}^4$ goes over into the operator of multiplication by the function $p \to \exp i\{x,p\}$, it follows that the operators Q_j are multiplication operators. A direct calculation shows that in the Hilbert space of U,

$$Q_0^2 - Q_1^2 - Q_2^2 - Q_3^2 = \lambda I, \tag{105}$$

where $\lambda = +m^2$ if U is associated with $X_m^{\pm}$ and $-m^2$ if it is associated with Y_m. If we now consider a function (scalar or vector valued) which belongs to the Hilbert space of U and vanishes outside a very small compact set, it follows that in the corresponding dynamical state, the momenta Q_1, Q_2, Q_3 and energy Q_0 have distributions with extremely small variance and hence we may approximately regard the particle as having sharply defined energy and momentum values. Now for a classical free relativistic particle, the quantity $Q_0{}^2-Q_1{}^2-Q_2{}^2-Q_3{}^2$ represents the square of its rest mass. Consequently, the constant λ must also be interpreted as the square of the rest mass. This rules out the alternative $\lambda=-m^2$ as otherwise it would seem that we have to admit particles with imaginary mass.

This leaves us with the representations associated with $X_m{}^{\pm}$ $(m\geq 0)$ and the trivial representation. The representations associated with $X_m{}^{\pm}$ represent free particles of rest mass m. The form of the representations suggests that at least for $m>0$ we interpret $U^{m,+,j}$ as representing a particle with spin j. We shall see much later (Section 7) that this is indeed the case. In the meantime we shall define the parameter j for the $U^{m,+,j}$ and the parameter $\frac{1}{2}n$ for the $U^{+,\pm n}$ to be the *spin* of the particle. The representation $U^{m,+,1/2}$ describes the *electron*; the representations $U^{m,+,0}$ represent the *mesons*. The representations $U^{+,\pm n}$, $U^{+,\rho}$ and $U^{+,\rho,'}$ represent the *massless* particles. We shall see later that $U^{+,\pm 1}$ represent the *neutrinos* and $U^{+,\pm 2}$ the *photons*.

The representations $U^{+,\rho}$ and $U^{+,\rho,'}$ correspond to *infinite dimensional* representations of the stability group E^*, and hence the states of the corresponding particles are represented by infinite dimensional vector valued functions on $X_0{}^{\pm}$. The physical interpretation of these particles is somewhat unclear, as they would have to describe particles whose spin observables have infinite spectra (cf. Section 5).

4. CLIFFORD ALGEBRAS

Clifford Algebras. The customary representation of the states of a relativistic electron is in terms of spinor fields in space-time. Such descriptions depend heavily on the properties of the Clifford algebra and its representations. We devote the present section to this topic.

Let V be an n-dimensional vector space over the complex field $\mathbf{C}$. We shall assume $n\geq 4$. Let B be a nonsingular symmetric bilinear form on $V\times V$. It is known that there exist bases $\{v_1, v_2,\cdots, v_n\}$ for V such that

$$B(v_r,v_s) = \begin{cases} 0, & r\neq s \\ 1, & r=s \end{cases}; \tag{106}$$

corresponding to such a basis we shall construct an associative algebra with unit element 1, which we shall call the *concrete Clifford algebra F*. We do this as follows. For each subset $A\subseteq N=\{1, 2,\cdots, n\}$ we choose an abstract element e_A with $e_\varnothing=1$, and define F to be the vector space over $\mathbf{C}$ with the e_A as a basis; $\dim(F)=2^n$. We shall first define a product operation between elements of the set e_A. To do this formally, we define, for A, $B\subseteq N$, $A\wedge B=A\cup B-A\cap B$; for $A\subseteq N$, $j\in N$, let $p(A,j)$ be the number of elements $i \in A$ with $i > j$; and for A, $B \subseteq N$, $p(A,B) = \sum_{j \in B} p(A,j)$. Let us then define

$$e_A e_B = (-1)^{p(A,B)} e_{A \wedge B}. \tag{107}$$

It is easy to prove that this operation is associative and that $1e_A = e_A 1 = e_A$ for all A. To prove the associativity, let A_1, A_2, $A_3 \subseteq N$. It is easily seen that $(A_1 \wedge A_2) \wedge A_3 = A_1 \wedge (A_2 \wedge A_3)$. On the other hand,

$$p(A_1, A_2 \wedge A_3) = \sum_{j \in A_2 \wedge A_3} p(A_1, j) \equiv p(A_1, A_2) + p(A_1, A_3) \pmod 2,$$

$$p(A_1 \wedge A_2, A_3) = \sum_{j \in A_3} p(A_1 \wedge A_2, j) \equiv p(A_1, A_3) + p(A_2, A_3) \pmod 2.$$

From these relations we obtain the relation

$$e_{A_1}(e_{A_2} e_{A_3}) = (e_{A_1} e_{A_2}) e_{A_3}$$

at once. Since the e_A span F, it follows that F becomes an associative algebra under the product operation defined by

$$\sum p_A e_A \sum q_A e_A = \sum (-1)^{p(A,B)} p_A q_B e_{A \wedge B}. \tag{108}$$

The mapping which sends v_r to $e_{\{r\}}$ extends to a linear isomorphism of V into F. Let us write τ for this linear map and write $e_r = \tau(v_r)$. From (107) we then obtain

$$e_r^2 = 1, \; e_r e_s + e_s e_r = 2\delta_{rs}, \tag{109}$$

and for $A = \{i_1, i_2, \cdots, i_r\}$ with $1 \leq i_1 < i_2 < \cdots < i_r \leq n$,

$$e_A = e_{i_1} e_{i_2} \cdots e_{i_r}. \tag{110}$$

Thus F is generated by the e_r and the generators satisfy the relations (109). From (109) we obtain the following equation:

$$\tau(v)^2 = B(v,v)1 \qquad (v \in V). \tag{111}$$

The construction of F which we have given here is not an invariant one since we have used the basis $\{v_1, \cdots, v_n\}$ explicitly. We shall now indicate a more abstract construction. Let $\mathfrak{T}$ be the tensor algebra over V and let $\mathfrak{J}$ be the two-sided ideal in $\mathfrak{T}$ generated by the set of all elements of the form

$$a_v = v \otimes v - B(v,v)1. \tag{112}$$

We observe that $\mathfrak{J}$ is *proper*. In fact, consider the concrete Clifford algebra F associated with a basis $\{v_1, \cdots, v_n\}$ satisfying (106). By the basic property of the tensor algebra $\mathfrak{T}$, there exists a homomorphism, $\bar{\tau}$ say, of $\mathfrak{T}$ into F, such that $\bar{\tau} = \tau$ on V. Since $\tau[V]$ generates F, $\bar{\tau}$ maps $\mathfrak{T}$ *onto* F. Since $\tau(v)^2 = B(v,v)1$ in F, $\bar{\tau}(a_v) = 0$ so that $\mathfrak{J} \subseteq \text{kernel}\,(\bar{\tau})$. This proves that $\mathfrak{J}$ is proper and, indeed, that $\dim \mathfrak{T}/\mathfrak{J} \geq 2^n$.

We shall now demonstrate that $\bar{\tau}$ has precisely $\mathfrak{J}$ as its kernel. This will be done if we show that dim $\mathfrak{T}/\mathfrak{J} \leq 2^n$. Consider the basis $\{v_1, \cdots, v_n\}$. Every element in $\mathfrak{T}$ is a linear combination of elements of the form $v_{i_1} \otimes \cdots \otimes v_{i_r}$. Now, $v_i \otimes v_i \equiv 1$ and $v_i \otimes v_j \equiv -v_j \otimes v_i$ modulo $\mathfrak{J}$, and hence it is obvious that $v_{i_1} \otimes \cdots \otimes v_{i_r} \equiv \pm w$ modulo $\mathfrak{J}$, where w is of the form $v_{j_1} \otimes \cdots \otimes v_{j_s}$ with $1 \leq j_1 < j_2 < \cdots < j_s \leq n$. For a given s, there are $\binom{n}{s}$ such tensors and hence dim $(\mathfrak{T}/\mathfrak{J}) \leq \binom{n}{0} + \binom{n}{1} + \cdots = 2^n$. This proves that $\mathfrak{J}$ is the kernel of $\bar{\tau}$.

Let π be the canonical homomorphism of $\mathfrak{T}$ onto $\mathfrak{T}/\mathfrak{J}$. We write $\mathscr{C}$ for $\mathfrak{T}/\mathfrak{J}$ and call it the *Clifford algebra* of V associated with B; when there is no doubt as to what B is, we omit the references to B. Since $\bar{\tau}$ is identical with τ on V and since τ is a linear isomorphism on V, it follows that π maps V isomorphically into $\mathscr{C}$.

The Clifford algebra $\mathscr{C}$ has the following *universal* property; if $\mathscr{C}'$ is an associative algebra with unit $1'$ and τ' is a linear map of V into $\mathscr{C}'$ such that (i) $\tau'[V]$ generates $\mathscr{C}'$, and (ii) $\tau'(v)^2 = B(v,v)1'$ for all $v \in V$, then there exists a homomorphism $\tau^{\sim}$ of $\mathscr{C}$ onto $\mathscr{C}'$ and one only, such that the diagram

(113)

$$\begin{tikzcd}
& \mathscr{C} \arrow[dd, "\tau^{\sim}"] \\
V \arrow[ur, "\pi"] \arrow[dr, "\tau'"'] & \\
& \mathscr{C}'
\end{tikzcd}$$

commutes; for there exists a homomorphism $\bar{\tau}'$ of $\mathfrak{T}$ into $\mathscr{C}'$ such that $\bar{\tau}' = \tau'$ on V, and as $\tau'[V]$ generates $\mathscr{C}'$, $\bar{\tau}'$ is surjective. By (ii) above, $a_v \in$ kernel $(\bar{\tau}')$ so that $\mathfrak{J} \subseteq$ kernel $(\bar{\tau}')$ and hence $\bar{\tau}'$ induces a unique homomorphism $\tau^{\sim}$ of $\mathscr{C}$ onto $\mathscr{C}'$ such that (113) commutes. Furthermore, this universal property determines the pair $(\mathscr{C},\pi)$ up to isomorphism. In fact, let $\mathscr{C}'$ be an associative algebra with unit $1'$ and π' a linear mapping of V into $\mathscr{C}'$ such that (i) $\pi'[v]$ generates $\mathscr{C}'$ and $\pi'(v)^2 = B(v,v)1'$ for all $v \in V$, (ii) $(\mathscr{C}',\pi')$ has the universal property mentioned above. Then there are homomorphisms θ and θ', such that θ maps $\mathscr{C}$ onto $\mathscr{C}'$, and θ' maps $\mathscr{C}'$ onto $\mathscr{C}$, and the diagram

(114)

$$\begin{tikzcd}
& \mathscr{C} \arrow[dd, shift right, "\theta"'] \\
V \arrow[ur, "\pi"] \arrow[dr, "\pi'"'] & \\
& \mathscr{C}' \arrow[uu, shift right, "\theta'"']
\end{tikzcd}$$

commutes. Diagram (114) shows that $\theta \circ \theta'$ and $\theta' \circ \theta$ are both identities and hence that θ and θ' are isomorphisms.

The invariant construction of $\mathscr{C}$ enables us to associate automorphisms of $\mathscr{C}$ with B-orthogonal transformations of V. More precisely, let L be an invertible endomorphism of V such that

$$B(Lv,Lv) = B(v,v)$$

for all v. Clearly, there exists an algebra automorphism $L^{\sim}$ of $\mathfrak{T}$ which extends L. It is obvious from (112) that $a_{Lv} = L^{\sim}(a_v)$ for all $v \in V$ from which we easily conclude that $L^{\sim}(\mathfrak{J}) = \mathfrak{J}$. This proves that $L^{\sim}$ induces an automorphism of $\mathscr{C}$. We write this automorphism as $x \to x^L$ $(x \in \mathscr{C})$. Since $L \to L^{\sim}$ is a homomorphism, we conclude that the mapping, which assigns to L the automorphism $x \to x^L$ of $\mathscr{C}$, is a homomorphism of the orthogonal group of B into the group of automorphisms of $\mathscr{C}$. Note that

$$\pi(Lv) = \pi(v)^L. \tag{115}$$

Simplicity. The central fact in the theory of Clifford algebras is that when $\dim V = n$ is *even*, $\mathscr{C}$ is simple. We shall now proceed to prove this result. Suppose that $\mathscr{D}$ is an arbitrary nonzero two-sided ideal in $\mathscr{C}$. Any nonzero element $u \in \mathscr{D}$ can be written uniquely as

$$u = p_0 1 + \sum_{1 \le i_1 < i_2 < \cdots < i_r \le n} p_{i_1,\cdots,i_r} \pi(v_{i_1}) \cdots \pi(v_{i_r}). \tag{116}$$

We shall choose u to be a nonzero element of $\mathscr{D}$ with the smallest number of nonzero $p_{i_1,\cdots,i_r}$ in its expansion (116). Now, a direct calculation shows that when r is odd, $\pi(v_{i_1}) \cdots \pi(v_{i_r})$ and $\pi(v_j)$ commute if j is one of $i_1, \cdots, i_r$ and anticommute if j is none of $i_1, \cdots, i_r$; whereas, when r is even, the situation is exactly the reverse. Consequently, when n itself is even and there is at least one $p_{i_1,\cdots,i_r} \neq 0$ in (116) above, one can find s, $(k_1, \cdots, k_s)$, and some j, such that $p_{k_1,\cdots,k_s} \neq 0$ and $\pi(v_{k_1}) \cdots \pi(v_{k_s})$ anticommutes with $\pi(v_j)$. As $\pi(v_j)[\pi(v_{i_1}) \cdots \pi(v_{i_r})]\pi(v_j) = \pm \pi(v_{i_1}) \cdots \pi(v_{i_r})$ any way, we obtain

$$\begin{aligned}\pi(v_j)u\pi(v_j) = p_0 1 + \sum_{(i_1,\cdots,i_r) \neq (k_1,\cdots,k_s)} p'_{i_1,\cdots,i_r} \pi(v_{i_1}) \cdots \pi(v_{i_r}) \\ - p_{k_1,\cdots,k_s} \pi(v_{k_1}) \cdots \pi(v_{k_s}),\end{aligned} \tag{117}$$

where $p'_{i_1,\cdots,i_r} = \pm p_{i_1,\cdots,i_r}$. Writing $u' = \frac{1}{2}(u + \pi(v_j)u\pi(v_j))$, we get, on noting that $u' \in \mathscr{D}$,

$$u' = p_0 1 + \sum_{(i_1,\cdots,i_r) \neq (k_1,\cdots,k_s)} p''_{i_1,\ldots,i_r} \pi(v_{i_1}) \cdots \pi(v_{i_r}),$$

where $p''_{i_1,\cdots,i_r} = \frac{1}{2}(p_{i_1,\cdots,i_r} + p'_{i_1,\cdots,i_r})$. Since the number of nonzero $p''_{i_1,\cdots,i_r}$ is strictly less than the number of nonzero $p_{i_1,\cdots,i_r}$, we have a contradiction to the minimal nature of u. This shows that all $p_{i_1,\cdots,i_r}$ must vanish in (116). But then $p_0 \neq 0$ and hence $1 \in \mathscr{D}$, showing that $\mathscr{D} = \mathscr{C}$. We have therefore proved that $\mathscr{C}$ is simple when n is even.

The Spin Representation. By the classical Wedderburn structure theorem (cf., for example, Weyl [2], pp. 79–96, 280–290), there exists, up to equivalence, exactly one finite dimensional irreducible representation of the simple associative algebra $\mathscr{C}$ when n is even. Let us select a representation of this unique equivalence class and denote it by ρ. It is known that in the left regular representation of $\mathscr{C}$, ρ occurs exactly as many times (as a direct summand) as its dimension. Hence

$$\dim \rho = 2^{n/2} \qquad (n \text{ even}). \tag{118}$$

Let us write Σ for the vector space in which ρ acts. Let G_B be the orthogonal group of V and, for any $L \in G_B$, let $x \to x^L$ be the automorphism of $\mathscr{C}$ induced by L. Then $x \to \rho(x^L)$ is also an irreducible representation of $\mathscr{C}$ in Σ. Consequently, there exists an invertible linear transformation S_1 of Σ such that

$$\rho(x^L) = S_1 \rho(x) S_1^{-1} \tag{119}$$

for all $x \in \mathscr{C}$; note that this is guaranteed as soon as we have

$$\rho(\pi(Lv)) = S_1 \rho(\pi(v)) S_1^{-1}$$

for all $v \in V$. We may even choose S_1 such that $\det S_1 = 1$. Now, by Schur's lemma, S_1 is determined by (119) up to a constant multiplying factor. Hence, if $L, L' \in G_B$ and we choose S_1, S_1', and S_0 such that (119) is satisfied for L, L', and LL' with S_1, S_1', and S_0, respectively, we obtain the relation

$$S_0 = \sigma S_1 S_1' \qquad (\sigma \in \mathbf{C}); \tag{120}$$

the determinant condition on S_0, S_1, and S_1' implies that $|\sigma| = 1$.

Let G_Σ be the lcsc group of all invertible linear transformations of Σ of determinant 1. The set of all $(L,S_1) \in G_B \times G_\Sigma$, such that

$$\rho(\pi(Lv))S_1 = S_1 \rho(\pi(v))$$

for all $v \in V$, is obviously a closed subset of $G_B \times G_\Sigma$, and the map $(L,S_1) \to L$ is continuous and maps this set *onto* G_B. Hence, by the Federer-Morse lemma (cf. Chapter V), there exists a Borel map $L \to S_1(L)$ of G_B into G_Σ such that (119) is satisfied by $S_1(L)$ for all $x \in \mathscr{C}$. But then (120) implies that $L \to S_1(L)$ is a projective (nonunitary) representation of G_B in Σ. We have

$$S_1(LL') = m(L,L')S_1(L)S_1(L')$$

for all $L, L' \in G_B$. The associative law $L(L'L'') = (LL')L''$ shows at once that m is a multiplier for G_B. Since G_B is semisimple, we may use the theory of multipliers (Chapter VII) to lift S_1 to a *representation* S' of the covering group G_B^* of G_B. As G_B^* does not have any one-dimensional representations other than the trivial one, S' is uniquely determined by S_1. $S'(\ell \to S'(\ell))$ is thus the unique representation of $G_B{}^*$ in Σ such that

$$\rho(\pi(L^\ell v)) = S'(\ell)\rho(\pi(v))S'(\ell)^{-1} \tag{121}$$

for all $v \in V$, $\ell \to L^\ell$ being the covering homomorphism of $G_B{}^*$ onto G_B.

The elements of Σ will be called *spinors* and the representation $S'(\ell \to S'(\ell))$ and its irreducible components are called the *spin representations*. We treat the elements of Σ as determining a new type of quantity, say q, such that when V undergoes the orthogonal transformation L^ℓ, q undergoes the transformation $S'(\ell)$. It was E. Cartan who discovered, in infinitesimal form, the spin representations; the global version was constructed by H. Weyl. Cartan also proved that the spin representations of $G_B{}^*$ *cannot* be obtained by decomposing the tensor algebra over V for $n \geq 4$ with respect to the natural action of $G_B{}^*$. The reader who is interested in a deeper study of these matters should consult Cartan's book [1], Chevalley's account [4], and the article of Brauer and Weyl [1].

We shall now extend these results to the case of real vector spaces and quadratic forms. Let V^0 be a real vector space of dimension n, n being even. Let B^0 be a nonsingular symmetric bilinear form on $V^0 \times V^0$. We write V for the complexification of V^0 and B for the bilinear form on $V \times V$ which extends B^0. As $n \geq 4$, the group of invertible endomorphisms of V^0 of determinant 1 which leave B^0 invariant, is semisimple. We write G_{B^0} for the connected component of the identity of this group. Identifying each endomorphism of V^0 with its natural extension to V (as a complex endomorphism), we have $G_{B^0} \subseteq G_B$. Let $G_{B^0}{}^*$ be the universal covering group of G_{B^0}. Then, denoting by Σ the space of all spinors, we have a projective representation

$$S_1 \colon L \to S_1(L)$$

of G_{B^0} in Σ such that $\rho(\pi(Lv)) = S_1(L)\rho(\pi(v))S_1(L)^{-1}$ for all $L \in G_{B^0}$ and $v \in V^0$. Since G_{B^0} is semisimple we may argue as before to deduce the existence of a unique representation

$$\ell \to S^0(\ell)$$

of $G_{B^0}{}^*$ in Σ such that

$$\rho(\pi(L^\ell v)) = S^0(\ell)\rho(\pi(v))S^0(\ell)^{-1} \tag{122}$$

for all $v \in V^0$ and $\ell \in G_{B^0}{}^*$, $\ell \to L^\ell$ being the covering map. It is usual to call S^0 and its components also as spin representations.

It is usual to exhibit S^0 in a particular way involving the so-called Dirac γ matrices. We shall select a basis $\{w_0, \cdots, w_{n-1}\}$ for V^0 such that

$$B^0(w_r, w_s) = \varepsilon_r \delta_{rs}, \tag{123}$$

where

$$\varepsilon_r = \begin{cases} +1 & (0 \le r \le p-1), \\ -1 & (p \le r \le n-1). \end{cases} \tag{124}$$

If F is the matrix $\operatorname{diag}\{\varepsilon_0, \varepsilon_1, \cdots, \varepsilon_{n-1}\}$, then G_{B^0} can be identified with the connected component of the identity of the group of all matrices L such that $L^tFL = F$. Since $F^{-1} = F$, we can argue exactly as we did in the Lorentz case to conclude that $L^t \in G_{B^0}$, whenever $L \in G_{B^0}$. Since $G_{B^0}{}^*$ is the universal covering group of G_{B^0}, there exists an anti-automorphism

$$\ell \to \ell^* \tag{125}$$

of $G_{B^0}{}^*$ such that it induces the anti-automorphism $L \to L^t$ of G_{B^0} (cf. (55)). If we write

$$S(\ell) = S^0(\ell^*)^{-1} \qquad (\ell \in G_{B^0}{}^*), \tag{126}$$

then it follows from (122) that $\ell \to S(\ell)$ is the unique representation of $G_{B^0}{}^*$ in Σ such that

$$S(\ell)^{-1}\rho(\pi(v))S(\ell) = \rho(\pi((L^{\ell})^t v)) \tag{127}$$

for all $v \in V^0$ and $\ell \in G_{B^0}{}^*$. Let us now write

$$\gamma_r = \rho(\pi(w_r)). \tag{128}$$

Then, as $\rho(\pi(v)^2) = B(v,v)1$ for all $v \in V^0$, we have

$$\begin{cases} \gamma_r{}^2 = \varepsilon_r 1 & (0 \le r \le n-1), \\ \gamma_r\gamma_s + \gamma_s\gamma_r = 0 & (r \ne s). \end{cases} \tag{129}$$

Conversely, given endomorphisms $\gamma_0, \cdots, \gamma_{n-1}$ of some vector space Σ' satisfying (129), there exists obviously a unique representation ρ' of the Clifford algebra $\mathscr{C}$ in Σ' such that $\rho'(\pi(w_r)) = \gamma_r$; in fact, we need only to define ρ' by this equation on V and extend it to $\mathscr{C}$ using the universal property (113). ρ' will then be equivalent to a multiple of ρ. Hence $\dim \Sigma'$ will be a multiple of $2^{n/2}$. If $\dim \Sigma' = 2^{n/2}$, then ρ' will itself be irreducible, leading to the irreducibility of the set $\{\gamma_0, \cdots, \gamma_{n-1}\}$. In this case, the equation (127) becomes

$$S(\ell)^{-1}\gamma_r S(\ell) = \sum_{s=0}^{n-1} a_{rs}(L^{\ell})\gamma_s, \tag{130}$$

where $a_{rs}(L)$ is the r-sth matrix entry of L with respect to the basis $\{w_0, \cdots, w_{n-1}\}$.

We summarize our discussion so far in this section in the form of the following theorem.

Theorem 9.5. *Let V be a complex n-space, $n \ge 4$, and even Let B be a nonsingular symmetric bilinear form on $V \times V$. Then, the Clifford algebra* *over V associated with B is simple, and has, up to (linear) equivalence, exactly one irreducible finite dimensional representation* ρ; $\dim \rho = 2^{n/2}$. *Suppose now that V^0 is a real linear subspace of V of real dimension n such that the restriction B^0, of B to $V^0 \times V^0$, is real. Let $\{w_0, \cdots, w_{n-1}\}$ be a basis for V^0 such that the relations* (123) *are satisfied. Then there exists a set $\{\gamma_0, \cdots, \gamma_{n-1}\}$ of $2^{n/2} \times 2^{n/2}$ complex matrices such that the relations* (129) *are satisfied. $\{\gamma_0, \gamma_1, \cdots, \gamma_{n-1}\}$ is necessarily irreducible and any two sets of such γ's are connected by a similarity transformation $\gamma_j \to A\gamma_j A^{-1}$. Finally, if $n \geq 4$, there exists a unique representation S of the universal covering group of the connected component of the identity of G_{B^0} such that the relations* (130) *are satisfied.*

The matrices γ_r are usually known as *Dirac's γ matrices.*

Consider now the case when $n=4$ and V^0 has a basis $\{w_0, w_1, w_2, w_3\}$ such that $B(w_0, w_0) = 1$, $B(w_r, w_r) = -1$ $(r = 1, 2, 3)$. We are then dealing with a Minkowskian form. We shall now exhibit a particular choice of the γ's. Let $\sigma_1, \sigma_2, \sigma_3$ be the matrices ((55) of Chapter VIII) defined by

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix},$$

and let $\gamma_0, \gamma_1, \gamma_2, \gamma_3$ be the 4×4 matrices defined by

$$\gamma_0 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \gamma_r = \begin{pmatrix} 0 & -\sigma_r \\ \sigma_r & 0 \end{pmatrix} \qquad (r = 1, 2, 3). \tag{131}$$

It is then a routine matter to check that the relations (129) are satisfied. Since the γ's are 4×4 matrices and their irreducible realization is also four-dimensional, $\{\gamma_0, \gamma_1, \gamma_2, \gamma_3\}$ must be irreducible. In this case, G_{B^0} is the connected homogeneous Lorentz group H, $G_B{}^*$ is the group $SL(2, \mathbf{C}) = H^*$, and the covering map is what we have been denoting by $m \to \delta(m)$. The relations (130) become:

$$S(m)^{-1}\gamma_r S(m) = \sum_{s=0}^{3} a_{rs}(\delta(m))\gamma_s \qquad (r = 0, 1, 2, 3). \tag{132}$$

γ_0 is a Hermitian matrix; γ_1, γ_2, and γ_3 are skew Hermitian.

We may use (132) to analyze the representation S of H^* in greater detail. Let $\mathfrak{h}^*$ (cf. (52)) be the Lie algebra of H^*. Then, we have, from (132),

$$[\dot{S}(X), \gamma_r] = -\sum_{s=0}^{3} a_{rs}(\dot{\delta}(X))\gamma_s \qquad (r = 0, 1, 2, 3) \tag{133}$$

(for $\dot{\delta}$, see (54)). We shall now determine the matrices $\dot{S}(X)$ in terms of the γ's. Since the γ's are irreducible, we can write $\dot{S}(X)$ as a linear combination of the 16 elements 1, γ_r, $\gamma_r\gamma_s$, $\gamma_r\gamma_s\gamma_t$, $\gamma_0\gamma_1\gamma_2\gamma_3$. The commutation rules

(129) and (133) imply easily that each $\dot{S}(X)$ must be a linear combination of 1 and the $\gamma_r\gamma_s$ $(r<s)$. Since the matrices of any representation of a semisimple Lie algebra are of trace 0,† and since $\mathrm{tr}(\gamma_r\gamma_s)=-\mathrm{tr}(\gamma_s\gamma_r)=0$ for $r\neq s$, $\dot{S}(X)$ must be a combination of only the $\gamma_r\gamma_s$. Let

$$\dot{S}(X) = \sum_{0\leq r<s\leq 3} c_{rs}(X)\gamma_r\gamma_s.$$

Substituting this into (133) we obtain, after a brief calculation,

$$c_{rs}(X) = \tfrac{1}{2}a_{sr}(\dot{\delta}(X))$$

so that

$$\dot{S}(X) = \tfrac{1}{2}\sum_{0\leq r<s\leq 3} a_{sr}(\dot{\delta}(X))\gamma_r\gamma_s. \tag{134}$$

If we now use the formula (54), we find, after some calculation,

$$\dot{S}(X) = \begin{pmatrix} X & 0 \\ 0 & -X^* \end{pmatrix}, \tag{135}$$

from which we obtain the formula

$$S(m) = \begin{pmatrix} m & 0 \\ 0 & (m^{-1})^* \end{pmatrix}; \tag{136}$$

here the asterisk denotes complex adjunction. Formulas (135) and (136) display the structure of the representation S explicitly.

The notion of a spinor field on any manifold on which the inhomogeneous Lorentz group acts can now be introduced. Let X be the space on which the inhomogeneous group $G=H\times' R^4$ acts. Let Σ be the space of four complex dimensions on which the spin representation S of H^* acts. A spinor field is then a function ψ from X into Σ with the following transformation property: if we transform X into itself by the element $(\delta(h),x)$ where x is the translation and $\delta(h)$ is the homogeneous component, then ψ transforms into the function ψ', where

$$\psi'(y) = S(h)\psi((\delta(h),x)^{-1}\cdot y) \qquad (y\in X). \tag{137}$$

5. REPRESENTATIONS IN VECTOR BUNDLES AND WAVE EQUATIONS

The descriptions obtained in Section 3 for the representations of the inhomogeneous Lorentz group are not quite suitable for making explicit some of the physical features of free relativistic particles. Also explicit calculations become difficult since the formulas of Section 3 involve certain cocycles whose existence was only proved abstractly. Our first concern in this section is to give more geometric descriptions of these representations which will be free of the disadvantages mentioned above. In each case we shall construct an appropriate vector bundle on the orbit in question, and exhibit the states of the particle as square integrable sections of this bundle. These results were first rigorously obtained by Bargmann and Wigner [1]. We shall follow their discussion without any essential modifications.

---

† Since any element of the Lie algebra is a sum of elements of the form $[X,Y]$ (cf. Chevalley [3], p. 67).

We shall next prove that these sections can be regarded as *tempered* distributions (in the sense of L. Schwartz) in momentum space $\mathbf{P}^4$. This enables us to take Fourier transforms and obtain the Hilbert spaces as spaces of (tempered) distributions on $\mathbf{R}^4$. The differential equations satisfied by these distributions will then be the wave equations. We shall see at this stage the famous Dirac and Maxwell equations making their appearance.

The distributions in space-time which describe the states of the particles are actually functions at least for a dense linear manifold of states. These functions are then seen to transform under the Lorentz group exactly like spinor fields, thus tying up the theory with the usual principles of relativistic physics according to which the states of the free particles are to be represented by spin fields in space-time.

The Representations $U^{m,+,j}$ Corresponding to Mass m, Spin j. We shall fix $m>0$ and consider the orbit $X_m{}^+$ with invariant measure $\alpha_m{}^+$. We introduce the complex 4-space $\mathbf{C}^4$ in which the Dirac matrices $\gamma_0, \gamma_1, \gamma_2, \gamma_3$ act as linear transformations, and treat the elements of $\mathbf{C}^4$ as spinors. The spin representation $h \to S(h)$ $(h \in H^*)$ is, of course, determined by the equations (132). We take for the γ's the matrices defined by the relation (131). We write $\{e_0,e_1,e_2,e_3\}$ for the standard basis vectors of $\mathbf{C}^4$; e_j has the components δ_{j0}, δ_{j1}, δ_{j2}, δ_{j3}. We write $\langle .,. \rangle$ for the usual *Hermitian* inner product for $\mathbf{C}^4$; $\langle e_r,e_s\rangle = \delta_{rs}$.

We begin by defining the bundle $B_m^{+,1/2}$ as

$$B_m^{+,1/2} = \left\{(p,v) : p \in X_m{}^+,\ v \in \mathbf{C}^4,\ \sum_{k=0}^{3} p_k\gamma_k v = mv\right\} \tag{138}$$

with the projection

$$\pi : (p,v) \to p. \tag{139}$$

An easy calculation based on (132) and (55) shows that if $(p,v) \in B_m^{+,1/2}$, then so does $(\delta(h)p,S(h^{*-1})v)$. Thus

$$h,(p,v) \to (p,v)^h = (\delta(h)p,S(h^{*-1})v) \tag{140}$$

converts $B_m^{+,1/2}$ into a H^*-space. The action is smooth and projects to the usual H^*-action on $X_m{}^+$. For any $p \in X_m{}^+$ the fiber at p is the linear subspace of $\mathbf{C}^4$ given by

$$B_m^{+,1/2}(p) = \left\{(p,v) : \sum_{k=0}^{3} p_k\gamma_k v = mv\right\}. \tag{141}$$

For the point $(m,0,0,0)$, (141) reduces to the equation

$$\gamma_0 v = v.$$

From our formula (131) for γ_0 we see the fiber at $(m,0,0,0)$ is the two-dimensional subspace spanned by e_1+e_3 and e_2+e_4. Hence (141) also defines a two-dimensional subspace.

We shall construct next a family of covariant inner products on the fibers of the bundle. From the formulas (136) for $S(h)$ we obtain

$$S(h)^*\gamma_0 S(h) = \gamma_0 \tag{142}$$

after a trivial calculation. Hence the representation S leaves invariant the *Hermitian* form $v \to m^{-1}\langle\gamma_0 v,v\rangle$ in $\mathbf{C}^4$. This form is actually positive definite on each fiber of the bundle $B_m^{+,1/2}$. To see this, we consider a $(p,v) \in B_m^{+,1/2}$. We have:

$$\sum_{k=0}^{3} p_k\langle\gamma_k v,v\rangle = m\langle v,v\rangle.$$

Now, $\langle\gamma_0 v,v\rangle$ is real as γ_0 is Hermitian, and $\langle\gamma_k v,v\rangle$ is purely imaginary, as γ_k is skew Hermitian. Hence, as the right side of the above equation is real, we must have $\sum_{k=1}^{3} p_k\langle\gamma_k v,v\rangle = 0$. We may therefore write

$$p_0\langle\gamma_0 v,v\rangle = m\langle v,v\rangle$$

or

$$m^{-1}\langle\gamma_0 v,v\rangle = p_0{}^{-1}\langle v,v\rangle \qquad (v \in B_m{}^{+,1/2}(p)). \tag{143}$$

Equation (143) exhibits explicitly the positive definite character of the form $v \to m^{-1}\langle\gamma_0 v,v\rangle$ on the fibers of $B_m^{+,1/2}$.

We now introduce the Hilbert space $\mathscr{H}^{+,1/2}$ of Borel sections of the bundle $B_m^{+,1/2}$. The value of $\|\varphi\|^2$ for a section $\varphi(p \to \varphi(p))$ is simply the integral

$$\int_{X_m^+} p_0{}^{-1}\langle\varphi(p),\varphi(p)\rangle d\alpha_m{}^+(p).$$

Now, using the formula in theorem 9.2 for $\alpha_m{}^+$, we can rewrite this as

$$\|\varphi\|^2 = \int_{\mathbf{P}^3} \langle\varphi(p),\varphi(p)\rangle d\beta_m^{+,1}(\mathbf{p}), \tag{144}$$

where

$$d\beta_m^{+,1}(\mathbf{p}) = \frac{dp_1dp_2dp_3}{2(p_1{}^2+p_2{}^2+p_3{}^2+m^2)}. \tag{145}$$

The representation U of G^* in $\mathscr{H}_m^{+,1/2}$ is given by

$$(U_{h,x}\varphi)(p) = \exp i\{x,p\}\varphi(\delta(h)^{-1}p)^h, \tag{146}$$

where $\varphi(p)^h$ is defined by (140).

To identify U with $U^{m,+,1/2}$ we use theorem 6.20. We must accordingly determine the representation of the stability group K^*, induced on the fiber at $(m,0,0,0)$. This fiber is the subspace of all vectors of $\mathbf{C}^4$ of the form $\begin{pmatrix} v \\ v \end{pmatrix}$ where $v \in \mathbf{C}^2$; and for $h \in K^*$, as $h = h^{*-1}$, (136) and (140) give us the action

$$\begin{pmatrix} v \\ v \end{pmatrix} \to \begin{pmatrix} hv \\ hv \end{pmatrix}.$$

This shows that the representation of K^* on the fiber at $(m,0,0,0)$ is just $D^{1/2}$. In view of theorem 6.20, we may conclude that the representation defined by (146) is equivalent to $U^{m,+,1/2}$.

We now proceed to the case of higher spin. We consider a half integer j $(\geq \frac{1}{2})$ and write $2j = N$. For each $\nu = 1, 2, \cdots, N$ we define $\gamma_r{}^\nu$ $(r = 0, 1, 2, 3)$ to be the operator in the N-fold tensor product $\mathbf{C}^4 \otimes \cdots \otimes \mathbf{C}^4 = \mathbf{C}^{4\otimes N}$ which sends $u_1 \otimes \cdots \otimes u_N$ to $u_1{}' \otimes \cdots \otimes u_N{}'$, where $u_i{}' = u_i$ for $i \neq \nu$ and $u_\nu{}' = \gamma_r u_\nu$: in the obvious notation,

$$\gamma_r{}^\nu = 1 \otimes \cdots \otimes \gamma_r \otimes 1 \otimes \cdots \otimes 1 \qquad (1 \leq \nu \leq N). \tag{147}$$

For $p \in X_m{}^+$ we define $B_m^{+,j}(p)$ as the set of all (p,t), where t belongs to the subspace of $\mathbf{C}^{4\otimes N}$ consisting of all *symmetric* tensors t such that

$$\sum_{r=0}^{3} p_r\gamma_r{}^\nu t = mt \qquad (\nu = 1, 2, \cdots, N). \tag{148}$$

We define $B_m^{+,j}$ by

$$B_m^{+,j} = \bigcup_{p \in X_m{}^+} B_m^{+,j}(p). \tag{149}$$

It is easy to verify that $B_m^{+,j}(p)$ is the subspace of the symmetric tensors which belong to $B_m^{+,1/2}(p) \otimes \cdots \otimes B_m^{+,1/2}(p)$.

For $(p,t) \in B_m^{+,j}$ and $h \in H^*$, we define

$$(p,t)^h = (\delta(h)p, S_N(h^{*-1})t), \tag{150}$$

where

$$S_N(h) = S(h) \otimes \cdots \otimes S(h). \tag{151}$$

$S_N(h^{*-1})$ maps the space $B_m^{+,1/2}(p) \otimes \cdots \otimes B_m^{+,1/2}(p)$ onto the space $B_m^{+,1/2}(\delta(h)p) \otimes \cdots \otimes B_m^{+,1/2}(\delta(h)p)$, and leaves the space of symmetric tensors invariant. Hence (150) converts $B_m^{+,j}$ into a H^*-space. Furthermore, S_N leaves invariant the Hermitian form

$$t \to m^{-N}\langle(\gamma_0 \otimes \cdots \otimes \gamma_0)t, t\rangle_N,$$

where $\langle .,. \rangle_N$ denotes the natural Hermitian inner product in $\mathbf{C}^{4\otimes N}$; (143) shows that for all $t \in B_m^{+,1/2}(p) \otimes \cdots \otimes B_m^{+,1/2}(p)$ one has

$$m^{-N}\langle(\gamma_0 \otimes \cdots \otimes \gamma_0)t, t\rangle_N = p_0{}^{-N}\langle t,t\rangle_N$$

so that this Hermitian form is positive definite on $B_m^{+,1/2}(p) \otimes \cdots \otimes B_m^{+,1/2}(p)$, and hence, in particular, on each fiber of $B_m^{+,j}$. The representation of the stability group K^* induced on the fiber at $(m,0,0,0)$ being the natural one carried by the space of symmetric tensors on the fiber of $B_m^{+,1/2}$ at $(m,0,0,0)$, is equivalent to D^j (cf. Chapter VIII). We now define $\mathscr{H}_m^{+,j}$ to be the Hilbert space of Borel sections φ of $B_m^{+,j}$ with

$$\|\varphi\|^2 = \int_{R^3} \langle\varphi(p),\varphi(p)\rangle_N d\beta_m^{+,N}(\mathbf{p}), \tag{152}$$

where

$$d\beta_m^{+,N}(\mathbf{p}) = \frac{dp_1dp_2dp_3}{2(p_1{}^2+p_2{}^2+p_3{}^2+m^2)^{(N+1)/2}}. \tag{153}$$

The representation U is defined by

$$(U_{h,x}\varphi)(p) = \exp i\{x,p\}\varphi(\delta(h)^{-1}p)^h. \tag{154}$$

$U \simeq U^{m,+,j}$ by theorem 6.20.

The reader might have noticed that we have not treated the spinless case in terms of the spinor calculus. This can also be done provided we consider, instead of the bundle (149), the bundle where the fibers are made up of *skew symmetric* tensors. More precisely, let

$$B_m^{+,0} = \{(p,t) : p \in X_m{}^+, t \in \mathbf{C}^4 \otimes \mathbf{C}^4, t \text{ skew symmetric}, \quad (\textstyle\sum p_r\gamma_r{}^\nu)t = mt, \nu = 1, 2\}, \tag{155}$$

with the norm defined for any Borel section φ of the bundle by

$$\|\varphi\|^2 = \int_{R^3} \langle\varphi(p),\varphi(p)\rangle_2 d\beta_m^{+,2}(\mathbf{p}). \tag{156}$$

The representation of G^* is then defined in the same way as (154). For a given $p \in X_m{}^+$, the fiber of $B_m^{+,0}$ at p consists of all skew symmetric elements of $B_m^{+,1/2}(p) \otimes B_m^{+,1/2}(p)$ which form an one-dimensional space. As the unitary group K^* has no nontrivial one-dimensional representations, it induces the trivial representation on the fiber of $B_m^{+,0}$ at $(m,0,0,0)$. Hence the representation of G^* defined in the Hilbert space of sections of $B_m^{+,0}$ is equivalent to $U^{m,+,0}$.

The Representations $U^{+,\pm n}$. These representations may be obtained by passing to the limit in $U^{m,+,j}$ as $m \to 0+$. We write

$$B_0{}^+ = \left\{(p,v) : p \in X_0{}^+,\ v \in \mathbf{C}^4,\quad \sum_{r=0}^{3} p_r\gamma_r v = 0\right\}. \tag{157}$$

$B_0{}^+$ is a bundle over $X_0{}^+$ and the action $h,(p,v) \to (p,v)^h$, where

$$(p,v)^h = (\delta(h)p, S(h^{*-1})v) \tag{158}$$

converts it into a H^*-space. For any $m > 0$, the fiber of $B_m^{+,1/2}$ at $p^{(m)} = ((1+m^2)^{1/2},0,0,1)$ is spanned by the vectors

$$v_1{}^{(m)} = \tfrac{1}{2}me_1 + \tfrac{1}{2}(1+(1+m^2)^{1/2})e_3,$$
$$v_2{}^{(m)} = \tfrac{1}{2}(1+(1+m^2)^{1/2})e_2 + \tfrac{1}{2}me_4,$$

which converge, as $m \to 0+$, to e_3 and e_2, respectively. Now e_3 and e_2 span the fiber of $B_0{}^+$ at $(1,0,0,1)$. Since H^* is transitive on $X_m{}^+$ and $X_0{}^+$, the same is true at any other point: if $p \in X_0{}^+$, there are points $p^{(m)} \in X_m{}^+$ which converge to p as $m \to 0+$, with the property that any vector v in the fiber of $B_0{}^+$ at p can be expressed as $\lim v^{(m)}$, where $v^{(m)}$ belongs to the fiber of $B_m^{+,1/2}$ at $p^{(m)}$. However, all this could be done with $-m$ instead of m, showing that the bundles $B_{-m}^{+,1/2}$ also collapse to $B_0{}^+$. On the other hand, the endomorphism

$$\Gamma = i\gamma_0\gamma_1\gamma_2\gamma_3 \tag{159}$$

transforms $B_m^{+,1/2}(p)$ into $B_{-m}^{+,1/2}(p)$ for each $p \in X_m{}^+$ for any $m > 0$, as is easily seen from the fact that Γ *anticommutes* with all the γ_r. In the limit, Γ leaves the fibers of $B_0{}^+$ invariant, leading to further degeneracies.

We calculate quickly that

$$\Gamma = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{160}$$

Hence Γ commutes with all $S(h)$ (cf. (136)). Thus, if $(p,v) \in B_0{}^+$, $(p,\Gamma v) \in B_0{}^+$. On the fiber at $(1,0,0,1)$ Γ has the eigenvalues ± 1. Hence the same is true on all the fibers. We may now define the bundles $B_0^{+,\pm 1/2}$ by

$$B_0^{+,\pm 1/2} = \{(p,v) : (p,v) \in B_0{}^+,\ \Gamma v = \mp v\}. \tag{161}$$

H^* leaves both $B_0^{+,+1/2}$ and $B_0^{+,-1/2}$ invariant. A simple calculation shows that for the stability group E^* at $(1,0,0,1)$ the representation induced on the fibers at $(1,0,0,1)$ becomes

$$m_{z,a} \to z^{\pm 1}. \tag{162}$$

We have already seen that the Hermitian form

$$v \to p_0{}^{-1}\langle v,v\rangle$$

is left invariant by the $S(h)$ and is positive definite on the fibers of $B_m^{+,1/2}$. Therefore, letting $m \to 0+$, we may conclude that the form is still invariant. At (1,0,0,1), the fiber of $B_0{}^+$ is spanned by e_3 and e_2 on which the form considered is even positive definite. Hence, it is positive definite on each fiber. We are thus in a position to apply theorem 6.20. $\mathscr{H}_0^{+,\pm 1/2}$ is the Hilbert space of Borel sections φ of $B_0^{+,\pm 1/2}$ with

$$\|\varphi\|^2 = \int_{\mathbf{P}^3} \langle\varphi(p),\varphi(p)\rangle d\beta_0^{+,1}(\mathbf{p}), \tag{163}$$

where

$$d\beta_0^{+,1}(\mathbf{p}) = \frac{dp_1dp_2dp_3}{2(p_1{}^2+p_2{}^2+p_3{}^2)}. \tag{164}$$

The representation U, which is equivalent to $U^{+,\pm 1}$, is defined by

$$(U_{h,x}\varphi)(p) = \exp i\{x,p\}\varphi(\delta(h)^{-1}p)^h. \tag{165}$$

For higher n, we proceed as in the case of nonzero mass. We write for $n \geq 1$,

$$B_0^{+,\pm n/2} = \left\{(p,t) : p \in X_0{}^+,\ t \in \mathbf{C}^{4\otimes n},\ \left(\sum_{r=0}^{3} p_r\gamma_r{}^\nu\right)t = 0,\right. \\ \left. \Gamma^\nu t = \mp t,\ \nu = 1, 2,\cdots, n\right\}, \tag{166}$$

where

$$\Gamma^\nu = i\gamma_0{}^\nu\gamma_1{}^\nu\gamma_2{}^\nu\gamma_3{}^\nu. \tag{167}$$

The action of H^* on these bundles is the obvious one. The norm of a section φ is defined by

$$\|\varphi\|^2 = \int_{\mathbf{P}^3} \langle\varphi(p),\varphi(p)\rangle_n d\beta_0^{+,n}(\mathbf{p}), \tag{168}$$

with

$$d\beta_0^{+,n}(\mathbf{p}) = \frac{dp_1dp_2dp_3}{2(p_1{}^2+p_2{}^2+p_3{}^2)^{(n+1)/2}}. \tag{169}$$

We shall end this discussion with a description of the representation $U^{+,2} \oplus U^{+,-2}$ which lends itself more readily to physical interpretation. We consider the bundle

$$B = \left\{(p,v) : p \in X_0{}^+,\ v \in \mathbf{C}^4,\ \{p,v\} = p_0v_0 - \sum_{r=1}^{3} p_rv_r = 0\right\}. \tag{170}$$

H^* acts on B as follows:

$$(p,v)^h = (\delta(h)p, \delta(h)v). \tag{171}$$

For $v, v' \in \mathbf{C}^4$ we introduce the *Lorentz-Hermite form*

$$Q(v,v') = -\{v,v'^*\} = -v_0v_0'^* + v_1v_1'^* + v_2v_2'^* + v_3v_3'^*.$$

We may regard the fiber $B(p)$ of B at p as the *complexified* tangent space to X_0 at p so that B is simply the complexified tangent bundle, and (171) gives the usual action on such bundles. Since $\delta(h)$ is real for each h, $\delta(h)$ preserves this Hermitian form also. We claim that $v \to -\{v,v^*\}$ is *nonnegative* on each fiber of B. It is enough to prove this at $(1,0,0,1)$, where the fiber is spanned by $w_1=(1,0,0,1)$, $w_2=(0,1,0,0)$, and $w_3=(0,0,1,0)$. The form, expressed in the basis just described, becomes $a_1w_1+a_2w_2+a_3w_3 \to |a_2|^2+|a_3|^2$. Let $\mathscr{H}^\sim$ be the space of equivalence classes of Borel sections φ of B for which

$$\|\varphi\|^2 = \int_{X_0^+} -\{\varphi(p),\varphi(p)^*\}d\alpha_0^+(p) < \infty. \tag{172}$$

$\mathscr{H}^\sim$ is complete, but $\|\cdot\|$ as defined in (172) is *not a norm, but only a seminorm.* In order that $\|\varphi\|=0$, it is necessary and sufficient that the vector $(\varphi_0(p),\varphi_1(p),\varphi_2(p),\varphi_3(p))$ be proportional to (p_0,p_1,p_2,p_3), i.e., the relations

$$p_j\varphi_k - p_k\varphi_j = 0 \qquad (0 \le j < k \le 3) \tag{173}$$

be satisfied for α_0^+-almost all p. We write $\mathscr{H}$ for the *Hilbert space* obtained by going over to the quotient space modulo (173).

The representation $U^\sim$ is first defined in $\mathscr{H}^\sim$ by setting

$$(U^\sim_{h,x}\varphi)(p) = \exp i\{x,p\}\varphi(\delta(h)^{-1}p)^h.$$

We now observe that if $\{\varphi(p),\varphi(p)^*\}=0$ for almost all p, the same is true for $U^\sim_{h,x}\varphi$ also. Hence $U^\sim$ induces a representation, say U, in $\mathscr{H}$. To compute the equivalence class of U we proceed, in view of theorem 6.20, to determine the representation of the stability group E^* at $(1,0,0,1)$, induced on the quotient space of the fiber of B at $(1,0,0,1)$, modulo the one-dimensional subspace spanned by $(1,0,0,1)$ itself. Since w_1, w_2, w_3 span the fiber at $(1,0,0,1)$ we find, modulo w_1, that the matrix $m_{z,a}$ acts as follows:

$$\begin{aligned} w_2 &\to \mathrm{Re}(z^2)w_2 - \mathrm{Im}(z^2)w_3, \\ w_3 &\to \mathrm{Im}(z^2)w_2 + \mathrm{Re}(z^2)w_3. \end{aligned} \tag{174}$$

Equations (174) show that the representation of E^* in question is the direct sum of the two characters

$$m_{z,a} \to z^{\pm 2}.$$

This proves that $U \simeq U^{+,2} \oplus U^{+,-2}$.

The Representations $U^{+,\rho}$ and $U^{+,\rho,'}$. These cannot be obtained by any limiting process from the representations corresponding to particles of nonzero rest mass. Since the inducing representations are infinite dimensional, these are not of much physical interest as the usual interpretation forces the spin observables to have infinite spectra. We do not discuss these in this book. The reader who is interested may refer to the articles of Bargmann-Wigner [1] and Wigner [5].

The Infinitesimal Operators. One of the advantages of describing the representations of the group G^* in Hilbert spaces of sections of various vector bundles is that the infinitesimal operators, which are in general unbounded, are all applicable to the *smooth* sections of the bundle. We can thus study the infinitesimal representation in a very explicit fashion. If X is any element of the Lie algebra of G^* and U is a representation of G^*, we write $X^{\sim}$ for the self-adjoint operator such that

$$\exp(itX^{\sim}) = U_{\exp tX}. \tag{175}$$

The calculations involved in computing the various operators X are very simple, and we confine ourselves to the determination of the angular momenta M_{rs}, $1 \leq r, s \leq 3$ (cf. Section 2 for the definition of the elements M_{rs}).

Mass $m > 0$, Spin j. We obtain from (134) the equation

$$\dot{S}_N(X) = \frac{1}{2}\sum_{\nu=1}^{N}\sum_{k<\ell} a_{\ell k}(\dot{\delta}(X))\gamma_k{}^{\nu}\gamma_\ell{}^{\nu}$$

for the differential of the representation (cf. (151)) S_N. Choosing X in the Lie algebra of H^* so that $\dot{\delta}(X) = M_{k\ell}$ $(1 \leq k < \ell \leq 3)$, we find, from (154),

$$M_{k\ell}{}^{\sim}\varphi = M_{k\ell;s}^{\sim}\varphi + M_{k\ell;0}^{\sim}\varphi, \tag{176}$$

where (φ is a smooth section with compact support)

$$M_{k\ell;s}^{\sim}\varphi = \left(\tfrac{1}{2}i\sum_{\nu=1}^{N}\gamma_k{}^{\nu}\gamma_\ell{}^{\nu}\right)\varphi \tag{177}$$

and

$$M_{k\ell;0}^{\sim}\varphi = i\left(p_\ell\frac{\partial}{\partial p_k} - p_k\frac{\partial}{\partial p_\ell}\right)\varphi. \tag{178}$$

In formulas (176) through (178), $1 \leq k < \ell \leq 3$. The operators $M_{k\ell;0}^{\sim}$ represent the *orbital angular momenta*, and the operators $M_{k\ell;s}^{\sim}$ represent the *spin angular momenta*. Note that the operators $M_{k\ell;s}^{\sim}$ $(k < \ell)$ are bounded, whereas the operators $M_{k\ell;0}^{\sim}$ are essentially self-adjoint when restricted to the linear manifold of the C^∞ sections with compact support of the bundle, as is easily seen by applying van Hove's lemma (cf. Chapter VIII) to the manifold $X_m{}^{+}$.

The case of the zero mass particles is discussed similarly. The formulas (176) through (178) remain valid without any change.

The Wave Equations. The traditional way of describing the states of a free particle is to put down a linear differential equation in space-time whose solutions will form the linear space, the rays of which are to describe the dynamical states of the particles. The equation will in general be invariant under the relativity group so that there will be a natural representation of the relativity group in the space of solutions. Such equations are known as *wave equations.* We shall now obtain wave equations corresponding to the various particles.

There are a few technical points to be attended to in the formulation of these equations. In the first place, the solutions in general will be Schwartz distributions in space-time. In the second place, not every solution will represent a state; the correct way to single out the relevant solutions is to define a norm on the space of solutions and take only those whose norm is finite. These norms, which are of course Hilbertian, are in general quite difficult to describe, because the general element is not a function but a distribution. The physical literature in general does not pay attention to these details. An a priori approach to the theory of such Hilbert spaces of distributions on manifolds has been given by Schwartz [3].

From our point of view, we overcome these difficulties quite simply because the Hilbert spaces are already precisely defined; the states are certain square integrable functions (scalar or vector valued) on the orbits. We then treat these as measures on $\mathbf{P}^4$ and take their Fourier transforms. The transformed space will consist of distributions, and we can obtain the form of the equation satisfied by them by a look at the measures on momentum space. It might be mentioned that when there is no momentum space (for example, when space-time is curved), the above approach will fail. The problem of rigorous descriptions of the wave equations in such cases must be considered open.

Given a distribution in $\mathbf{P}^4$, one cannot define its Fourier transform unless it is *tempered.* We recall, therefore, the definition of tempered distributions (cf. Schwartz [2]). Let V be a real n-dimensional Euclidean space with the positive definite product $\langle .,. \rangle$ and Lebesgue measure dv. For any translation invariant differential operator D and any complex polynomial q on V, the function

$$\varphi \to \sup_{v \in V} |q(v)(D\varphi)(v)|$$

is a seminorm on $C_c^\infty(V)$ and the collection of these seminorms induces a locally convex topology for $C_c^\infty(V)$; its completion may be identified with $\mathscr{S}(V)$, the space of C^∞ functions on V for which

$$\sup_{v \in V} |q(v)(D\varphi)(v)| < \infty$$

for all q and D. If $\{v_1, \cdots, v_n\}$ is an orthonormal basis for V, the operators $D_{\nu_1,\cdots,\nu_n} = (\partial/\partial x_1)^{\nu_1} \cdots (\partial/\partial x_n)^{\nu_n}$ linearly span the algebra of all translation invariant differential operators, $x_1, \cdots, x_n$ being the linear coordinate functions on V associated with the chosen basis. The topology can then be also induced by the collection of the seminorms

$$\sup_{x_1,\cdots,x_n} |(1+x_1^2+\cdots+x_n^2)^k (D_{\nu_1,\cdots,\nu_n}\varphi)(x_1,\cdots,x_n)|$$

for various $k=0, 1, 2, \cdots$ and $\nu_1, \cdots, \nu_n \geq 0$. A tempered distribution E is a complex valued linear functional on $C_c^\infty(V)$ which is continuous in the topology induced by these collections of seminorms. By extending these to $\mathscr{S}(V)$ we may also regard a tempered distribution as a continuous linear functional on $\mathscr{S}(V)$.

For any $\varphi \in \mathscr{S}(V)$, the function $\hat{\varphi}$, where

$$\hat{\varphi}(x) = (2\pi)^{-n/2} \int_V \exp[-i\langle x,v\rangle]\varphi(v)dv \qquad (x \in V)$$

is also in $\mathscr{S}(V)$ and

$$\varphi \to \hat{\varphi}$$

is an automorphism of $\mathscr{S}(V)$. Given any tempered distribution

$$E(\varphi \to E(\varphi),\ \varphi \in \mathscr{S}(V)),$$

we may then define its Fourier transform $\hat{E}$ by the equation

$$\hat{E}(\varphi) = E(\hat{\varphi}). \tag{179}$$

If L is any linear automorphism of V of determinant ± 1, L induces, in a natural fashion, an automorphism of the space $\mathscr{S}(V)$, and (by duality) of the space of tempered distributions. We write

$$(E^L)(\varphi) = E(\varphi^{L^{-1}}), \tag{180}$$

where $\varphi^{L^{-1}}$ is the function $v \to \varphi(Lv)$. Taking Fourier transforms of (180), we get

$$(E^L)^\wedge = \hat{E}^{(L^0)}, \tag{181}$$

where L^0 is defined by

$$\langle L^0 v, v'\rangle = \langle v, L^{-1}v'\rangle \qquad (v, v' \in V). \tag{182}$$

The map

$$\mathscr{F} : E \to \hat{E}$$

is then a linear automorphism of the space of tempered distributions. Let $\{v_1, \cdots, v_n\}$ be an orthonormal basis, P and Q two complex polynomials of n real variables, and write

$$q(x_1v_1 + \cdots + x_nv_n) = Q(x_1, \cdots, x_n),$$
$$D_P = P(\partial/\partial x_1, \cdots, \partial/\partial x_n).$$

Then, for any tempered distribution E,

$$(qD_P E)\hat{} = D_{Q^*} P^* \hat{E},$$

where P^* is the polynomial

$$P^*(x_1 v_1 + \cdots + x_n v_n) = P(ix_1, \cdots, ix_n)$$

and

$$D_{Q^*} = Q\left(i\frac{\partial}{\partial x_1}, \cdots, i\frac{\partial}{\partial x_n}\right).$$

All this is well known and may be found in Schwartz's book [2].

The possibility of describing the free particles by relativistically invariant wave equations in $\mathbf{R}^4$ depends on the following lemma.

Lemma 9.6. *Let s be a real number >0 and f a complex valued Borel function on $\mathbf{P}^4$ such that*

$$\int_{X_m^+} |f(p)|^2 p_0^{-s} d\alpha_m^+(p) < \infty \qquad (m \geq 0). \tag{183}$$

Then the distribution†

$$T_f : \varphi \to \int_{X_m^+} f(p)\varphi(p) d\alpha_m^+(p) \qquad (\varphi \in C_c^\infty(\mathbf{P}^4))$$

is well defined, tempered, and its Fourier transform $\hat{T}_f$ satisfies the "wave equation"

$$(\Box + m^2)\hat{T}_f = 0, \tag{184}$$

where

$$\Box = \frac{\partial^2}{\partial x_0^2} - \frac{\partial^2}{\partial x_1^2} - \frac{\partial^2}{\partial x_2^2} - \frac{\partial^2}{\partial x_3^2}.$$

Proof. We shall show that T_f is a tempered distribution. Consider first the case $m > 0$. Then X_m^+ is a closed submanifold in $\mathbf{P}^4$ and the above integral obviously exists for all $\varphi \in C_c(\mathbf{P}^4)$. To prove that T_f is a tempered distribution, we take a $k > \frac{1}{2}s + 2$ and rewrite $T_f(\varphi)$ as

$$T_f(\varphi) = \frac{1}{2}\int_{\mathbf{P}^3} [f(p_0, \mathbf{p}) p_0^{-(s+1)/2}] G(\mathbf{p}) \psi(p_0, \mathbf{p}) d\mathbf{p},$$

where

$$G(\mathbf{p}) = p_0^{(s-1)/2}(1 + p_0^2 + p_1^2 + p_2^2 + p_3^2)^{-k/2},$$

$$\psi(p_0, \mathbf{p}) = \varphi(p_0, \mathbf{p})(1 + p_0^2 + p_1^2 + p_2^2 + p_3^2)^{k/2},$$

and

$$p_0 = (m^2 + p_1^2 + p_2^2 + p_3^2)^{1/2}.$$

---

† Actually a complex Borel measure.

Since both G and the function $\mathbf{p} \to f(p_0,\mathbf{p})p_0^{-(s+1)/2}$ are in $\mathscr{L}^2(d\mathbf{p})$, we have, for a suitable constant C, and all $\varphi \in C_c^{\infty}(P^4)$,

$$|T_f(\varphi)| \leq C \sup_{p \in P^4} |(1+\langle p,p\rangle)^k \varphi(p)|.$$

This proves at once that T_f is a tempered distribution.

We proceed to the case $m=0$. We shall first show that $T_f(\varphi)$ is well defined, i.e., that the function

$$\mathbf{p} \to f(p_0,\mathbf{p})\varphi(p_0,\mathbf{p})2p_0^{-1}$$

is Lebesgue integrable in the domain $p_0>0$ of $\mathbf{P}^3$. But as

$$\mathbf{p} \to f(p_0,\mathbf{p})p_0^{-(s+1)/2} \qquad \text{and} \qquad \mathbf{p} \to \varphi(p_0,\mathbf{p})p_0^{(s-1)/2}$$

are both square integrable in $\mathbf{P}^3$, the assertion follows. We now argue as before to conclude that

$$|T_f(\varphi)| \leq C \sup_{p \in P^4} |(1+\langle p,p\rangle)^k \varphi(p)|$$

for suitable constants C and k. This completes the proof that T_f is a tempered distribution.

We may therefore form the Fourier transform $\hat{T}_f$ of T_f. Since $T_f(\varphi)=0$ whenever $\varphi = 0$ on X_m^+, we have, for all $\varphi \in C_c^{\infty}(\mathbf{P}^4)$,

$$T_f((p_0^2-p_1^2-p_2^2-p_3^2-m^2)\varphi) = 0$$

or

$$(p_0^2-p_1^2-p_2^2-p_3^2-m^2)T_f = 0. \tag{185}$$

The distribution $(p_0^2-p_1^2-p_2^2-p_3^2-m^2)T_f$ is tempered and its Fourier transform is $(-\Box-m^2)\hat{T}_f$. Hence (185) leads to the equation

$$(\Box+m^2)\hat{T}_f = 0.$$

Corollary 9.7. *If f moreover vanishes outside a compact subset of* $\mathbf{P}^4$ *and is also bounded, then $\hat{T}_f$ is a function on* $\mathbf{R}^4$*—in fact, an entire function and the equation* (184) *is satisfied in the usual sense.*

Proof. T_f is now a finite complex measure on $\mathbf{P}^4$ with compact support. It is well known that $\hat{T}_f$ is then an entire function. Equation (184) then implies the same relation in the usual classical sense.

We are now in a position to describe the wave equations in $\mathbf{R}^4$. Let us first consider the case of mass m and spin $\frac{1}{2}$. The Hilbert space of the representation $U^{m,+,1/2}$ is the space of sections f of the bundle $B_m^{+,1/2}$ with the norm defined by (144). Writing

$$f = (f_0,f_1,f_2,f_3),$$

we have

$$\sum_{r=0}^{3} \gamma_r p_r f(p) = mf(p) \tag{186}$$

for $\alpha_m{}^+$-almost all p; the norm is given by

$$\|f\|^2 = \sum_{r=0}^{3} \int_{X_m{}^+} |f_r(p)|^2 p_0{}^{-1} d\alpha_m{}^+(p).$$

From lemma 9.6 we may now conclude that T_{f_r} is a tempered distribution on $\mathbf{P}^4$ for all $r=0, 1, 2, 3$. Let Φ be the Fourier transform of T_f, considered as a ($\mathbf{C}^4$-valued) vectorial distribution on $\mathbf{R}^4$,

$$\Phi = (\hat{T}_{f_0}, \hat{T}_{f_1}, \hat{T}_{f_2}, \hat{T}_{f_3}).$$

Then Φ is tempered, and we have, as a consequence of (186),

$$\sum_{r=0}^{3} i\gamma_r \frac{\partial}{\partial x_r} \Phi = m\Phi. \tag{187}$$

We observe that the equation

$$(\Box + m^2)\Phi = 0$$

follows from (187). In fact, the square of the operator $\sum_{r=0}^{3} i\gamma_r(\partial/\partial x_r)$ is, in view of the commutation rules (129), simply the operator $-\Box$; applying $\sum_{r=0}^{3} i\gamma_r(\partial/\partial x_r)$ to both sides of (187) we obtain (184). Equation (187) is called the *Dirac wave equation of the electron.*

The states of a free relativistic particle of rest mass $m>0$ and spin $\frac{1}{2}$ may thus be regarded as the rays of a certain Hilbert space of tempered $\mathbf{C}^4$-valued solutions of (187). If we write out the action of G^* on Φ using formulas (146) and (181), we see that Φ transforms like a spinor field on space-time (with respect to a suitable action of the inhomogeneous Lorentz group G on R^4). The equation (187) and the fact that the solutions Φ transform like spinor fields were first discovered by Dirac [2], and they constituted the first great step in the mathematical treatment of relativistic systems.

Using formulas (152), (154), and lemma 12.6, we can write the wave equations corresponding to arbitrary spin and arbitrary rest mass m. In the general case, with $m>0$, Φ is a tempered distribution on $\mathbf{R}^4$ whose values are tensors of rank N over $\mathbf{C}^4$. The wave equations are

$$\sum_{r=0}^{3} i\gamma_r{}^{\nu} \frac{\partial}{\partial x_r} \Phi = m\Phi \qquad (\nu = 1, 2, \cdots, N). \tag{188}$$

If we require the distribution Φ to take the values in the space of symmetric tensors of rank N over $\mathbf{C}^4$, then (188) corresponds to spin $\frac{1}{2}N$; as before, (188) implies the wave equation

$$(\Box + m^2)\Phi = 0.$$

If $N=2$ and the values of Φ are skew symmetric tensors, then (188) defines the spinless particle of nonzero mass. The zero mass equations are described similarly. The wave equations are

$$\sum_{r=0}^{3} i\gamma_r^{\nu} \frac{\partial}{\partial x_r}\Phi = 0 \qquad (\nu = 1, 2, \cdots, n) \tag{189}$$

and

$$i\gamma_0^{\nu}\gamma_1^{\nu}\gamma_2^{\nu}\gamma_3^{\nu}\Phi = -\Phi \qquad (\nu = 1, 2, \cdots, n) \tag{190}$$

for the spin $\frac{1}{2}n$ case; for spin $-\frac{1}{2}n$, the equation (190) is replaced by

$$i\gamma_0^{\nu}\gamma_1^{\nu}\gamma_2^{\nu}\gamma_3^{\nu}\Phi = \Phi \qquad (\nu = 1, 2, \cdots, n). \tag{191}$$

In all cases, therefore, we have obtained linear differential equations of the *first order* to describe the states of the particles. The early attempts to formulate relativistic one-particle theories led to the *second-order* equation $(\square + m^2)\Phi = 0$ and were therefore regarded with serious misgiving until Dirac discovered the equation (187).

It is sometimes asserted that the concept of spin enters inevitably in any relativistic description of a free particle as soon as one attempts to formulate Lorentz invariant first-order wave equations to describe the states of the particle. This is a misnomer inasmuch as even the spinless particles are described by such equations (cf. (155)). There is only one mathematically meaningful way to look at spin—namely, as a consequence of the method by which one differentiates between different systems which are covariant with respect to the group of motions of configuration space.

We have been somewhat vague as to which solutions of these wave equations are to be singled out for describing the quantum mechanical states of the particle. These solutions will have to be tempered, but this alone is not sufficient. In fact, consider a tempered vectorial solution Φ of (187). We may then write $\Phi = \hat{\Psi}$, where Ψ is a tempered distribution on $\mathbf{P}^4$. Since $(\square + m^2)\Phi = 0$, $(p_0^2 - p_1^2 - p_2^2 - p_3^2 - m^2)\Psi = 0$, so that Ψ vanishes for all elements of $C_c^{\infty}(\mathbf{P}^4)$ of the form

$$(p_0^2 - p_1^2 - p_2^2 - p_3^2 - m^2)\varphi \qquad \varphi \in C_c^{\infty}(\mathbf{P}^4).$$

It can be shown easily that an element of $C_c^{\infty}(\mathbf{P}^4)$ has this form if and only if it vanishes on $X_m^{+} \cup X_m^{-}$. Therefore, we can construct distributions Ψ^{+} and Ψ^{-} *living on the submanifolds* X_m^{+} and X_m^{-}, respectively, such that $\Psi = \Psi^{+} + \Psi^{-}$. Clearly, $\Psi^{-} = 0$ is a necessary condition (the so-called restriction to positive energy solutions). But the actual conditions are very complicated; these imply that $\Psi^{-} = 0$, Ψ^{+} is a measure on X_m^{+} absolutely continuous with respect to α_m^{+}, and $(d\Psi^{+}/d\alpha_m^{+}) \cdot p_0^{-1/2} \in$

$\mathscr{L}^2(X_m{}^+, \mathbf{C}^4, \alpha_m{}^+)$. It seems difficult to exhibit natural conditions on Φ itself which are equivalent to these.

We shall conclude this section with a few remarks on the representation $U^{+,2} \oplus U^{+,-2}$ which was realized in the space of sections of the complexified tangent bundle to $X_0{}^+$ (modulo the subspace of sections whose values at each $p \in X_0{}^+$ are proportional to p). If we write

$$f = (f_0, f_1, f_2, f_3) \tag{192}$$

for a section of the complexified tangent bundle of $X_0{}^+$, then (170) leads to

$$p_0 f_0(p) - \sum_{k=1}^{3} p_k f_k(p) = 0. \tag{193}$$

Let

$$\Phi = \hat{T}_f = (\Phi_0, \Phi_1, \Phi_2, \Phi_3),$$

where T_f is the vectorial measure,

$$T_f : \varphi \to \int_{X_0{}^+} f \varphi \, d\alpha_0{}^+.$$

Then we have, as $(p_0{}^2 - p_1{}^2 - p_2{}^2 - p_3{}^2) T_f = 0$,

$$\Box \Phi = 0.$$

Moreover, (193) gives us

$$\frac{\partial}{\partial x_0} \Phi_0 - \frac{\partial}{\partial x_1} \Phi_1 - \frac{\partial}{\partial x_2} \Phi_2 - \frac{\partial}{\partial x_3} \Phi_3 = 0. \tag{194}$$

However, different Φ will correspond in general to the same state; the condition that Φ and Φ' may represent the same vector in the Hilbert space of states is the Fourier transform of (173), i.e.,

$$\frac{\partial}{\partial x_k} \Psi_r - \frac{\partial}{\partial x_r} \Psi_k = 0 \qquad (r, k = 0, 1, 2, 3) \quad (\Phi - \Phi' = \Psi)$$

or

$$\operatorname{curl} \Psi = 0. \tag{195}$$

The equations $\Box \Phi = 0$ are the well known Maxwell equations; the condition (194) is the so-called *Lorentz* condition and (195) gives the *gauge condition*. In particle physics, the representation which we are discussing now is assumed to represent the *photon*. A closer analysis reveals that the two representations $U^{+,\pm 2}$ represent the photon in its two states of left and right circular polarization. We do not go into these ideas here.

6. INVARIANCE UNDER THE INVERSIONS

We have so far restricted our attention to the transformation properties of the states of a free particle under the *connected* inhomogeneous Lorentz group. We shall now examine the behavior of the states under the inversions also. The main result of this section asserts that the irreducible representation of the connected group G^* associated with a particle of mass $m>0$ can be enlarged in an essentially unique fashion to a representation of the complete inhomogeneous Lorentz group $\tilde{G}$ into the group of automorphisms of the states. We shall also obtain explicit formulas for the operators which correspond to the inversions. We show further that the irreducible representations $U^{+,\pm n}$ corresponding to zero mass and spin $\frac{1}{2}n$ do *not* possess this property of extension.

Consider now a representation

$$\alpha : (L,x) \to \alpha_{(L,x)}$$

of the complete inhomogeneous group $\tilde{G}$ into the group of automorphisms of the logic of a Hilbert space. We have seen already that the restriction of α to G is induced by a unique representation U of G^*. Let us now take an inversion I_τ ($\tau = s$, t, or st) and consider a symmetry S_τ of the Hilbert space $\mathscr{H}$ (on which α and U act) which induces $\alpha_{(I_\tau,0)}$. Since

$$\alpha_{(I_\tau L I_\tau, I_\tau x)} = \alpha_{(I_\tau,0)}\alpha_{(L,x)}\alpha_{(I_\tau,0)},$$

it follows that the representations $(h,x) \to S_\tau U_{(h,x)} S_\tau^{-1}$ and $(h,x) \to U_{(h_\tau, I_\tau x)}$ induce the same homomorphism into the projective group of $\mathscr{H}$ (cf. (69) for definition of h_τ). This means, as G^* has no nontrivial one dimensional representations, that

$$S_\tau U_{(h,x)} S_\tau^{-1} = U_{(h_\tau, I_\tau x)} \tag{196}$$

for all $(h,x) \in G^*$. Conversely, let $U(h,x \to U_{(h,x)})$ be a representation of G^* in a Hilbert space $\mathscr{H}$ whose logic is denoted by $\mathscr{L}$, and let $\alpha(L,x \to \alpha_{(L,x)})$ be a representation of G in Aut($\mathscr{L}$) such that α is induced by U. Since the group $\tilde{G}$ is the semidirect product of G and the group of inversions H_{inv}, it is obvious that for α to extend to a representation of $\tilde{G}$ into Aut($\mathscr{L}$), it is necessary and sufficient that there exist elements α_s, α_t, α_{st} of Aut($\mathscr{L}$) such that

$$\alpha_\tau^2 \quad \text{is the identity} \tag{197}$$

for $\tau = s,t,st$,

$$\alpha_{st} = \alpha_s \alpha_t \tag{198}$$

and, such that, if S_τ is a symmetry which induces α_τ, (196) is satisfied for all $(h,x) \in G^*$. The conditions (197) and (198) are the conditions for

$I_\tau \rightarrow \alpha_\tau$ to be a representation of H_{inv}, and (196) is just the condition characteristic of semidirect products. If such α_τ exist, then the map

$$(LI_\tau, x) \rightarrow \alpha_{(L,x)}\alpha_\tau \tag{199}$$

is a representation of $\tilde{G}$ which extends α from G.

We shall now analyze the conditions (196), (197), and (198). In the first place, S_τ can be unitary or anti-unitary and is determined by $\alpha_{(I_\tau,0)}$ up to a constant multiplying factor. Since I_τ is a reflection, S_τ^2 must be a multiple of the identity, i.e., $S_\tau^2 = \zeta 1$ for a constant ζ with $|\zeta| = 1$. If S_τ is unitary, we may, by replacing S_τ by $\zeta^{-1/2}S_\tau$, assume that $S_\tau^2 = 1$. On the other hand, for anti-unitary S_τ, $(cS_\tau)(cS_\tau) = S_\tau^2$ for all complex c with absolute value 1, so that in this case ζ is an intrinsic invariant of $\alpha_{(I_\tau,0)}$ itself. Actually, $\zeta = \pm 1$, as the following lemma shows.

Lemma 9.8. *Let S be an anti-unitary operator in a Hilbert space $\mathscr{H}$ such that $S^2 = \zeta 1$ for some complex constant ζ of modulus 1. Then $\zeta = \pm 1$.*

Proof. Let $\{e_j\}$ be an orthonormal basis for $\mathscr{H}$ and let $e_j' = Se_j$. Then $\{e_j'\}$ is an orthonormal basis also and there exists a unitary operator S_1 such that

$$S_1 e_j = e_j'$$

for all j. Write $S_2 = SS_1^{-1}$. Then S_2 is anti-unitary and $S_2^2 = 1$. From $S^2 = \zeta 1$ we obtain the equation $S_2 S_1 S_2^{-1} = \zeta S_1^{-1}$. If $E \rightarrow P_E$ is the spectral measure of S_1, based on the unit circle, then the last equation leads to

$$S_2 P_E S_2 = P_{\zeta E}$$

for all Borel subsets E of the unit circle. Since $S_2^2 = 1$, we obtain

$$P_E = P_{\zeta^2 E}$$

for all E. This means that $\zeta^2 = 1$.

Let U be a representation of G^* in a Hilbert space $\mathscr{H}$. Then there exists a unique projection valued measure Q on P^4 such that

$$U_{(1,x)} = \int_{P^4} \exp i\{x,p\} dQ(p)$$

for all $x \in R^4$. The support of Q is called the *spectrum* of U.

Lemma 9.9. *Let α be a representation of $\tilde{G}$, the complete inhomogeneous Lorentz group, into the group* Aut $\mathscr{L}$, *$\mathscr{L}$ being the logic of the Hilbert space $\mathscr{H}$. Let U be the representation of G^* which induces α on G. Suppose that* (a) *the restriction of U to the translation subgroup is not trivial,* (b) *there exists a constant c such that the spectrum of U is contained in the set $\{p : p_0 \geq c\}$. Then $\alpha_{(I_\tau,0)}$ is induced by an anti-unitary operator for $\tau = t$ or st, and by a unitary operator if $\tau = s$.*

Proof. Consider first time inversion. Let S_t be a symmetry of $\mathscr{H}$ which induces $\alpha_{(I_t,0)}$. Let M be the spectrum of U. Then M is a closed subset of P^4 invariant under H. Now on any orbit Y_m $(m>0)$, p_0 can take arbitrary negative values. So $M \subseteq \{p : p_0 \geq 0\}$. The condition (196) gives us

$$S_t U_{(1,x)} S_t^{-1} = U_{(1,I_t x)}.$$

Suppose now that S_t is unitary. Let Q be the projection valued measure of the representation $x \to U_{(1,x)}$. Then

$$\begin{aligned} U_{(1,I_t x)} &= \int_{\mathbf{P}^4} \exp i\{I_t x, p\} dQ(p) \\ &= \int_{\mathbf{P}^4} \exp i\{x, I_t p\} dQ(p), \end{aligned}$$

which shows that $x \to U_{(1,I_t x)}$ has the associated projection valued measure $E \to Q_{I_t[E]}$. Since S_t is unitary, these two projection valued measures would be equivalent. In particular, their spectra will be the same. This implies that the spectrum of Q is invariant under I_t. This implies that $M = \{0\}$. This contradicts the assumption (a) of the lemma and shows that $(I_t,0)$ must be represented by an anti-unitary operator. The argument for $(I_{st},0)$ is similar. Since $I_s = I_t I_{st}$, $(I_s,0)$ must be represented by a unitary operator.

We are now in a position to prove the main result of this section.

Theorem 9.10. *Let U be a representation of G^* in a Hilbert space $\mathscr{H}$ with the property that for some $m \geq 0$, $Q_{X_m^+}$ is the identity operator of $\mathscr{H}$, Q being the projection valued measure of the representation $x \to U_{1,x}$. Suppose U induces the representation α of the group G into* Aut($\mathscr{L}$), *$\mathscr{L}$ being the logic of $\mathscr{H}$. Then, in order that α may extend to a representation of the full inhomogeneous group $\tilde{G}$ into* Aut($\mathscr{L}$), *it is necessary that the representation of the stability group at some point of X_m^+, which is associated with U, be self-contragredient. If U is equivalent to one of $U^{m,+,j}$ $(m>0, j=0, \frac{1}{2}, 1, \cdots)$, $U^{+,0}$, or $U^{+,n} \oplus U^{+,-n}$ $(n=1, 2, \cdots)$, then U induces a representation α of G in* Aut($\mathscr{L}$), *and α has an extension to a representation of $\tilde{G}$ in* Aut($\mathscr{L}$); *in the first two cases, the extension is even unique.*

Proof. Suppose that U induces a representation α of G in Aut($\mathscr{L}$), and that α has an extension to $\tilde{G}$ as a representation into Aut($\mathscr{L}$). By lemma 12.9, $(I_\tau,0)$ must be represented by an anti-unitary operator S_τ when $\tau = st$. We have:

$$S_{st} U_{h,x} S_{st}^{-1} = U_{h,-x}$$

for all $(h,x) \in G^*$. Let p^0 be any point of X_m^+ and let π be a representation of the stability group H^0 at p^0 such that U is associated with X_m^+ and π. Then, a direct calculation shows that the representation $h, x \to U_{h,-x}$

is associated with $-X_m^+$ and π. On the other hand, as S is anti-unitary, $h, x \to U_{h,-x}$ is contragredient to U, and hence, by lemma 9.3, it is associated with $-X_m^+$ and π^0, where π^0 is contragredient to π. Therefore π and π^0 must be equivalent.

Let now U be equivalent to one of $U^{m,+,j}$ $(m>0)$, $U^{+,0}$, $U^{+,n} \oplus U^{+,-n}$. It is obvious that U maps the elements $(\pm 1,0)$ of G^* into multiples of the identity. Hence U induces a representation α of G into $\operatorname{Aut}(\mathscr{L})$. We shall first prove that if $U \simeq U^{m,+,j}$ $(m>0)$ or if $U \simeq U^{+,0}$, α has at most one extension to $\tilde{G}$ as a representation into $\operatorname{Aut}(\mathscr{L})$. It is enough to prove that for each τ, the element $\alpha_{(I_\tau,0)}$ of $\operatorname{Aut}(\mathscr{L})$ is uniquely determined by U. Suppose that $S_\tau^{(1)}$ and $S_\tau^{(2)}$ are two symmetries such that

$$S_\tau^{(r)} U_{h,x} (S_\tau^{(r)})^{-1} = U_{h_\tau, I_\tau x}$$

for all $h, x \in G^*$ and $r = 1, 2$. Then $(S_\tau^{(2)})^{-1} S_\tau^{(1)}$ commutes with all $U_{h,x}$ and, being a unitary operator,† it must be a multiple of the identity, as U is irreducible. Hence the corresponding elements of $\operatorname{Aut}(\mathscr{L})$ coincide.

We now come to the existence. A simple and direct argument can be given, but in view of the general usefulness of explicit formulas, we shall prove the existence part of the theorem by actually writing down the relevant symmetries. For $U = U^{m,+,0}$ we take $\mathscr{H} = \mathscr{L}^2(X_m^+, \alpha_m^+)$ and write the representation U in the form

$$(U_{h,x}\varphi)(p) = \exp i\{x,p\}\varphi(\delta(h)^{-1}p).$$

We define the element $\alpha_{(I_\tau,0)}$ of the group $\operatorname{Aut}(\mathscr{L})$ as the automorphism induced by the symmetry S_τ, where

$$\begin{aligned} (S_t\varphi)(p_0, \mathbf{p}) &= J\varphi(p_0, -\mathbf{p}), \\ (S_{st}\varphi)(p_0, \mathbf{p}) &= J\varphi(p_0, \mathbf{p}), \\ (S_s\varphi)(p_0, \mathbf{p}) &= \varphi(p_0, -\mathbf{p}), \end{aligned} \tag{200}$$

J being complex conjugation. It is easily checked that

$$S_{st} = S_s S_t$$

and that

$$S_\tau U_{h,x} S_\tau^{-1} = U_{h_\tau, I_\tau x}.$$

Note that these formulas also remain valid when $m = 0$.

For $U = U^{m,+,j}$ $(j = \frac{1}{2}, 1, \cdots)$ we use the vector bundle description (154). $\mathscr{H}$ is now the Hilbert space of square integrable sections of the vector bundle $B_m^{+,j}$. Let $N = 2j$ and let J_N be the natural conjugation in

---

† By lemma 12.9, $S_\tau^{(r)}$ is anti-unitary for $\tau = t, st$ and unitary for $\tau = s$.

$\mathbf{C} \otimes \cdots \otimes \mathbf{C}$ (N factors). For any endomorphism A in C let $A^{(N)}$ be the endomorphism $A \otimes \cdots \otimes A$ in $\mathbf{C} \otimes \cdots \otimes \mathbf{C}$. We now define S_τ by

$$\begin{aligned}(S_{st}\varphi)(p) &= (\gamma_0\gamma_1\gamma_3)^{(N)}J_N\varphi(p),\\ (S_t\varphi)(p) &= (\gamma_1\gamma_3)^{(N)}J_N\varphi(-I_tp),\\ (S_s\varphi)(p) &= \gamma_0{}^{(N)}\varphi(I_sp).\end{aligned} \tag{201}$$

Here, $p \in X_m{}^+$; note that $-I_tp = I_sp \in X_m{}^+$.

We observe that γ_0, γ_1, $i\gamma_2$, γ_3 are all real. Hence (cf. (147)) for $\nu = 1, \cdots, N$, $\gamma_0{}^\nu$, $\gamma_1{}^\nu$, $i\gamma_2{}^\nu$, $\gamma_3{}^\nu$ are also real. Therefore $\gamma_0{}^\nu$, $\gamma_1{}^\nu$, and $\gamma_3{}^\nu$ commute with J_N while $\gamma_2{}^\nu$ anticommutes with J_N. It also follows from the same observation that $\gamma_0^{(N)}$, $(\gamma_1\gamma_3)^{(N)}$, and $(\gamma_0\gamma_1\gamma_3)^{(N)}$ commute with J_N. This implies at once that

$$\begin{aligned}\left(\sum_{r=0}^{3} p_r\gamma_r{}^\nu\right)(\gamma_0\gamma_1\gamma_3)^{(N)}J_N &= (\gamma_0\gamma_1\gamma_3)^{(N)}J_N\left(\sum_{r=0}^{3} p_r\gamma_r{}^\nu\right),\\ \left(\sum_{r=0}^{3} p_r\gamma_r{}^\nu\right)(\gamma_1\gamma_3)^{(N)}J_N &= (\gamma_1\gamma_3)^{(N)}J_N\left(\sum_{r=0}^{3} p_r'\gamma_r{}^\nu\right),\\ \left(\sum_{r=0}^{3} p_r\gamma_r{}^\nu\right)\gamma_0^{(N)} &= \gamma_0^{(N)}\left(\sum_{r=0}^{3} p_r'\gamma_r{}^\nu\right),\end{aligned} \tag{202}$$

where $1 \leq \nu \leq N$, $p_0' = p_0$, $p_r' = -p_r$, $r = 1, 2, 3$. These equations show that the transformations S_τ map sections of $B_m^{+,j}$ into sections of $B_m^{+,j}$. Moreover, an easy calculation shows that $\gamma_r^{(N)}$ are unitary in $\mathbf{C} \otimes \cdots \otimes \mathbf{C}$. So we can use (152) to conclude that $\|S_\tau\varphi\|^2 = \|\varphi\|^2$. In other words, the S_τ are symmetries of $\mathscr{H}$. Also,

$$\begin{aligned}S_s{}^2 &= 1,\\ S_\tau{}^2 &= (-1)^N 1 \qquad (\tau = t, st),\\ S_{st} &= S_sS_t = S_tS_s.\end{aligned} \tag{203}$$

It follows from the equation $S(h)\gamma_0 = \gamma_0 S(h^{*-1})$ (cf. (136)) that S_s satisfies (196). We claim now that S_t also satisfies (196). This reduces to proving that

$$(\gamma_1\gamma_3)^{(N)}J_NS_N(h^{*-1}) = S_N(h)(\gamma_1\gamma_3)^{(N)}J_N. \tag{204}$$

Since J_N commutes with $(\gamma_1\gamma_3)^{(N)}$, this reduces to proving that

$$(\gamma_1\gamma_3)^{(N)}S(h^{*-1})^{(N)} = J_NS(h)^{(N)}(\gamma_1\gamma_3)^{(N)}J_N.$$

Now, J_N is the natural conjugation and so we have

$$J_NS(h)^{(N)}(\gamma_1\gamma_3)^{(N)}J_N = (JS(h)(\gamma_1\gamma_3)J)^{(N)}$$

where we write J for J_1. A trivial calculation shows that

$$(\gamma_1\gamma_3)S(h^{*-1}) = JS(h)(\gamma_1\gamma_3)J$$

and this proves (204).

The case $U \simeq U^{+,N} \oplus U^{+,-N}$ still remains to be examined. The symmetries S_τ are again defined by (201). They still satisfy (203). This time, the S_τ operate on the space of square integrable sections of the sum $B^{+,N/2} \oplus B^{+,-N/2}$. A trivial calculation (cf. (167)) shows that Γ^ν anticommutes with both $\gamma_0^{(N)}$ and $(\gamma_0\gamma_1\gamma_3)^{(N)}J_N$ while it commutes with $(\gamma_1\gamma_3)^{(N)}J_N$ $(\nu=1,\cdots,N)$. From this the results concerning U follow quickly. Note that both S_s and S_{st} intertwine $U^{+,N}$ and $U^{+,-N}$ while S_t leaves the subspaces corresponding to $U^{+,\pm N}$ invariant.

Corollary 9.11. *The representations $U^{+,\pm N}$ of G cannot be extended to representations of $\tilde{G}$ into the corresponding groups of automorphisms of the underlying logics.*

Proof. The representations of the stability subgroup at (1,0,0,1) are the characters $m_{z,a} \to z^{\pm N}$ which are not self-contragredient.

7. LOCALIZATION

In Chapter VIII we described up to physical equivalence all systems which are irreducibly covariant with respect to the group of all motions of a Euclidean configuration space. In three dimensions this led to the classification of these systems through their spin and to the familiar expressions for the position and momentum observables. On the other hand, in this chapter we have described the Lorentz invariant free particles and it would be a natural question to ask whether we could write expressions for the position operators of these particles. To make this question precise, we select a particular inertial coordinate system, and ask, for $x_0=0$ in this frame, whether a suitable projection valued measure $E \to P_E$ can be defined on the Borel sets of R^3, the projections acting on the Hilbert space $\mathscr{H}$ of the particle, such that the range of P_E represents the experimental statement that the particle at time $t=0$ belongs to E. We shall take the same point of view as we did in Chapter VIII and require $P(E \to P_E)$ to be such that the system is covariant with respect to the group of rigid motions of R^3. This condition can be rephrased in a simple manner. Let O be an observer and let x_0, x_1, x_2, x_3 be his time and space coordinates. We identify the space of O with R^3 and obtain the projection valued measure P on the σ-algebra of Borel subsets of R^3. If $(h,a) \in G^*$ is such that $a_0=0$ and $h \in K^*$, then the new observer O' determined by (h,a) differs from O only in space origin and space orientation. Since $U_{(h,a)}$ represents the operator which expresses the transformations of states associated with the change from O to O', the above mentioned condition of covariance can be expressed as follows:

$$(205) \qquad U_{h,a}P_E U_{h,a}^{-1} = P_{\delta(h)[E]+a} \qquad (h \in K^*, a_0 = 0).$$

The problem of describing the position operators can then be regarded as the problem of determining, for a given representation U of G^*, all the projection valued measures P such that (205) is satisfied.

Let us write V for the representation U when restricted to the subgroup of all (h,a) such that $a_0=0$ and $h \in K^*$. Let us call this subgroup M. The pair (V,P) is a system of imprimitivity for M based on $\mathbf{R}^3$ and hence we can use the results of Chapter VIII to conclude that there exists a unitary isomorphism W of $\mathscr{H}$ onto $\mathscr{L}^2(\mathbf{R}^3,\mathscr{K},d\mathbf{x})$ and a representation $g \to s(g)$ of the group K^* in $\mathscr{K}$ such that for $(h,a) \in M$,

$$(WV_{h,a}W^{-1}f)(\mathbf{x}) = s(h)f(\delta(h)^{-1}(\mathbf{x}-\mathbf{a})), \tag{206}$$

and

$$(WP_EW^{-1}f)(\mathbf{x}) = \chi_E(\mathbf{x})f(\mathbf{x}). \tag{207}$$

In other words, a necessary and sufficient condition that there should exist at least one projection valued measure P satisfying (205) is that the representation V be unitarily equivalent to the representation, say L^s, of M described by the right side of (206). If W is a unitary operator effecting this equivalence, P may then be defined by (207).

The representation appearing on the right side of (206) can be described in another, equivalent, form through the use of Fourier-Plancherel transforms. Let us consider the unique unitary automorphism

$$\mathscr{F} : f \to \hat{f} \tag{208}$$

of $\mathscr{L}^2(\mathbf{R}^3,\mathscr{K},d\mathbf{x})$ onto itself such that for all $\varphi \in C_c(\mathbf{R}^3)$,

$$\int_{\mathbf{R}^3} \hat{f}(\mathbf{x})\varphi(\mathbf{x})d\mathbf{x} = \int_{\mathbf{R}^3} f(\mathbf{x})\left(\int_{\mathbf{R}^3} (2\pi)^{-3/2}e^{-i\langle \mathbf{x},\mathbf{y}\rangle}\varphi(\mathbf{y})d\mathbf{y}\right)d\mathbf{x}. \tag{209}$$

Then a straightforward calculation shows that

$$(\mathscr{F}L^s_{h,a}\mathscr{F}^{-1}\hat{f})(\mathbf{x}) = \exp[-i\langle \mathbf{a},\mathbf{x}\rangle]s(h)\hat{f}(\delta(h)^{-1}\mathbf{x}), \tag{210}$$

where $\langle .,. \rangle$ is the usual Euclidean inner product. We may obviously change the measure $d\mathbf{x}$ to any rotation invariant measure in the same measure class without changing the form or the equivalence class of the representation defined by (210). We have therefore obtained the following lemma:

Lemma 9.12. *Let V be a representation of M in some Hilbert space $\mathscr{H}$. Let m be a real number ≥ 0. Then, for the existence of a projection valued measure P based on $\mathbf{R}^3$ such that (V,P) is a system of imprimitivity, it is necessary and sufficient that there exist a representation s,*

$$s : h \to s(h)$$

of K^* in a Hilbert space $\mathscr{K}$ such that V is equivalent to the representation V^s of M in $\mathscr{L}^2(\mathbf{R}^3,\mathscr{K},\alpha_m)$ defined by

$$(V^s_{h,a}f)(\mathbf{y}) = \exp[-i\langle \mathbf{a},\mathbf{y}\rangle]s(h)f(\delta(h)^{-1}\mathbf{y}), \tag{211}$$

α_m being the measure given by

$$d\alpha_m = \frac{d\mathbf{y}}{2(m^2+\langle \mathbf{y},\mathbf{y}\rangle)^{1/2}}. \tag{212}$$

Note that m plays a completely harmless role in this lemma.

We shall now discuss the problem of recognizing whether a given representation of M is equivalent to some V^s as defined by (211). To this end, we need to consider representations more general than the representations V^s. Let $X = \mathbf{R}^3 - \{\mathbf{0}\}$, and let us regard X as a K^*-space with the action h, $\mathbf{y} \to \delta(h)\mathbf{y}$. Given a strict (K^*,X)-cocycle C with values in the unitary group of a Hilbert space $\mathscr{K}$, we consider the representation V^C of M in $\mathscr{L}^2(X,\mathscr{K},\alpha_m)$ given by

$$(V^C_{h,a}f)(\mathbf{y}) = \exp[-i\langle \mathbf{a},\mathbf{y}\rangle]C(h,\delta(h)^{-1}\mathbf{y})f(\delta(h)^{-1}\mathbf{y}). \tag{213}$$

The representation V^s is obtained if we take

$$C(h,\mathbf{y}) = s(h).$$

The strict cocycle C determines for each $r>0$, the strict (K^*,S^1)-cocycle C_r,

$$C_r(h,\mathbf{u}) = C(h,r\mathbf{u}), \tag{214}$$

where S^1 is the unit sphere. Note that K^* is transitive on S^1.

The next lemma determines the equivalence class of the representation V^C completely and it is our basic tool in the subsequent analysis of localizability. We make a few comments on (213). The projection valued measure $Q(E \to Q_E)$, where Q_E is multiplication by χ_E, is determined in a direct fashion by the $V^C_{1,a}$:

$$V^C_{1,a} = \int_X \exp[-i\langle \mathbf{a},\mathbf{y}\rangle]dQ(\mathbf{y}).$$

From this it follows that the unitary equivalence class of Q is determined by the unitary equivalence class of V^C. In particular, the dimension of $\mathscr{K}$, which equals the multiplicity of Q, is determined by the unitary equivalence class of V^C.

We recall that T' is the subgroup of K^* consisting of all diagonal matrices

$$m_{\zeta,0} = \begin{pmatrix} \zeta & 0 \\ 0 & \zeta^{-1} \end{pmatrix} \qquad (|\zeta| = 1).$$

T' is the stability subgroup of K^* at the points $(0,0,r)$ $(r>0)$.

Lemma 9.13. *Let C^1 and C^2 be two strict (K^*,X)-cocycles with values in the unitary groups of Hilbert spaces $\mathscr{H}^1$ and $\mathscr{H}^2$. Then, the representations V^{C^1} and V^{C^2} of M are equivalent if and only if for almost all $r > 0$, the strict (K^*,S^1)-cocycles $C_r{}^1$ and $C_r{}^2$ define equivalent representations of the subgroup T' of K^*.*

Proof. Let $R^+ = \{r : r > 0\}$ and $Y = R^+ \times S^1$. The map $(r,\mathbf{u}) \rightarrow r\mathbf{u}$ is a homeomorphism of Y with X which sends $(r,\delta(h)\mathbf{u})$ to $\delta(h)(r\mathbf{u})$. It is easily shown that the measure α_m on X goes over to a multiple of $\beta \times \sigma$ on Y under this map, where σ is the normalized rotation invariant measure on S^1, and $d\beta = r^2(m^2+r^2)^{-1/2}dr$.

Suppose now that V^{C^1} and V^{C^2} are equivalent representations of the group M. Then $\dim \mathscr{H}^1 = \dim \mathscr{H}^2$ so that we may assume that $\mathscr{H}^1 = \mathscr{H}^2 = \mathscr{H}$. Let D be a unitary operator which effects the equivalence. D commutes with all the multiplication operators $A_{\mathbf{a}}$,

$$(A_{\mathbf{a}}f)(\mathbf{x}) = \exp[-i\langle \mathbf{a},\mathbf{x}\rangle]f(\mathbf{x}).$$

From this it follows easily that D commutes with all operators Q_E of the form

$$(Q_E f)(\mathbf{x}) = \chi_E(\mathbf{x})f(\mathbf{x}).$$

From lemma 6.4 we conclude that $D = D^{\sim}$ for some Borel map $x \rightarrow D(x)$ of X into the unitary group of $\mathscr{H}$, i.e.,

$$(Df)(\mathbf{x}) = D(\mathbf{x})f(\mathbf{x}) \qquad (\mathbf{x} \in X).$$

The condition that D intertwines the restrictions of V^{C^1} and V^{C^2} to K^* now leads to the equation

$$C^2(h,\mathbf{x}) = D(\delta(h)\mathbf{x})C^1(h,\mathbf{x})D(\mathbf{x})^{-1}$$

for each $h \in K^*$ for almost all $\mathbf{x}$. Going over to $Y = R^+ \times S^1$ and using Fubini theorem, we obtain from this the equation

$$C_r{}^2(h,\mathbf{u}) = D_r(\delta(h)\mathbf{u})C_r{}^1(h,\mathbf{u})D_r(\mathbf{u})^{-1}$$

for almost all triples $(h,r,\mathbf{u})$, where

$$D_r(\mathbf{u}) = D(r\mathbf{u}) \qquad (\mathbf{u} \in S^1).$$

In other words, for β-almost all r, the strict (K^*,S_1)-cocycles $C_r{}^1$ and $C_r{}^2$ are cohomologous. From theorem 5.27, as K^* is transitive on S_1, it now follows that for almost all r, the representations of the stability subgroup T' of K^*, defined by $C_r{}^1$ and $C_r{}^2$, are equivalent.

We now come to the converse. Suppose C^1 and C^2 satisfy the condition described above. We may assume that $\mathscr{H}^1 = \mathscr{H}^2 = \mathscr{H}$. Then the strict (K^*,S^1)-cocycles $C_r{}^1$ and $C_r{}^2$ are cohomologous for β-almost all r. Let

$$\mathscr{H}_1 = \mathscr{L}^2(S^1,\mathscr{H},\sigma).$$

Let the representations of K^* in $\mathscr{L}^2(S^1,\mathscr{K},\sigma)$—say $A^{1,r}$ and $A^{2,r}$—be defined by

$$(A^{j,r}_h f)(\mathbf{u}) = C_r^{\,j}(h,\delta(h)^{-1}\mathbf{u})f(\delta(h)^{-1}\mathbf{u}).$$

Let Q^1 be the projection valued measure based on S^1 defined by

$$(Q_E{}^1 f)(\mathbf{u}) = \chi_E(\mathbf{u})f(\mathbf{u}) \qquad (E \subseteq S^1).$$

Then, for any r for which $C_r{}^1$ and $C_r{}^2$ are cohomologous, there exists a unitary operator D_r in $\mathscr{K}_1$ commuting with all $Q_E{}^1$ and having the property (cf. theorem 6.12)

$$D_r \circ A^{1,r} \circ D_r{}^{-1} = A^{2,r}.$$

Since D_r commutes with Q^1, D_r commutes with the operator in $\mathscr{K}_1$ of multiplication by the function $\mathbf{u} \rightarrow \exp ir\langle\mathbf{a},\mathbf{u}\rangle$ for all $\mathbf{a}$, r. If we then write $W^{j,r}_{a,h}$ for the operator in $\mathscr{K}_1$ defined by

(215) $$(W^{j,r}_{h,a}f)(\mathbf{u}) = \exp[-ir\langle\mathbf{a},\mathbf{u}\rangle](A^{j,r}_h f)(u),$$

then

(216) $$D_r \circ W^{1,r}_{h,a} \circ D_r{}^{-1} = W^{2,r}_{h,a}$$

for all $(h,a) \in M$. $W^{1,r}$ and $W^{2,r}$ are representations of M in $\mathscr{K}_1$ and their dependence on r is Borel; (216) is valid for β-almost all $r > 0$.

We shall now show that we may choose D_r so that $r \rightarrow D_r$ is a Borel function. To this end, we rewrite (216) as follows:

$$D_r W^{1,r}_{h,a} = W^{2,r}_{h,a} D_r.$$

Let us now denote by $\mathscr{U}_1$ the unitary group of $\mathscr{K}_1$ and consider the set $\mathscr{D}$ of all pairs $(D,r) \in \mathscr{U}_1 \times R^+$ such that

$$D W^{1,r}_{h,a} = W^{2,r}_{h,a} D$$

for all $(h,a) \in M$. Since $W^{j,r}$ is a representation, it is a continuous map from M into $\mathscr{U}_1$ and hence it is clear that the set $\mathscr{D}$ is a Borel subset of $\mathscr{U}_1 \times R^+$. Equation (216) shows that the map $D, r \rightarrow r$ maps $\mathscr{D}$ onto almost all of R^+. It follows now from von Neumann's cross-section lemma (cf. Chapter V) that there exists a Borel set $N \subseteq R^+$ of measure zero and a Borel map

$$D' : r \rightarrow D'(r)$$

of $R^+ - N$ into $\mathscr{U}_1$ such that $(D'(r),r) \in \mathscr{D}$ for all $r \in R^+ - N$. Let $D'^{\sim}$ be the unique unitary operator in $\mathscr{L}^2(R^+,\mathscr{K}_1,\beta)$ defined by

$$(D'^{\sim} F)(r) = D'(r)F(r)$$

for $r \in R^+ - N$.

The definition of $W^{j,r}$ makes it obvious that the operators $W^j_{h,a}$ defined in $\mathscr{L}^2(R^+,\mathscr{K}_1,\beta)$ by

$$(W^j_{h,a}F)(r) = W^{j,r}_{h,a}F(r)$$

are unitary, that W^j is a representation of M, and that $D'^{\sim}$ intertwines W^1 and W^2. In particular, W^1 and W^2 are equivalent. Now, the Hilbert space $\mathscr{L}^2(R^+,\mathscr{K}_1,\beta)$ is canonically isomorphic to $\mathscr{L}^2(R^+ \times S^1, \mathscr{K}, \beta \times \sigma)$, which in turn is canonically isomorphic to $\mathscr{L}^2(X,\mathscr{K},\alpha_m)$ under the unitary isomorphism induced by the measure preserving transformation $r, \mathbf{u} \rightarrow r\mathbf{u}$ of $R^+ \times S^1$ onto X. A little calculation now shows that the representation W^j goes over to the representation V^{C^j} in the Hilbert space $\mathscr{L}^2(X,\mathscr{K},\alpha_m)$. Consequently, V^{C^1} and V^{C^2} are equivalent. This completes the proof of the lemma.

Corollary 9.14. *In order that there exist a projection-valued measure P based on $\mathbf{R}^3$ such that (V^{C^1},P) is a system of imprimitivity for the group M, it is necessary and sufficient that there exist a representation s*

$$s : h \rightarrow s(h)$$

of K^* acting in $\mathscr{K}$, such that for almost all $r>0$, the restriction of s to the subgroup T' is equivalent to the representation $h \rightarrow C^1(h,(0,0,r))$ of T'.

Proof. Let s be a representation of K^* and let C^s be the strict (K^*,X)-cocycle defined by

$$C^s : h, \mathbf{x} \rightarrow s(h).$$

Then the necessary and sufficient condition for the existence of a suitable P is simply that, for a suitable s, C^1 and C^s satisfy the condition described in lemma 9.13. The corollary now follows at once if we observe that $\delta(h)$ fixes $(0,0,r)$ if and only if $h = m_{\zeta,0}$ for a suitable ζ with $|\zeta| = 1$.

Corollary 9.15. *Let (V,P^0) be a system of imprimitivity based on $\mathbf{R}^3$ for the group M and let P^0 have finite multiplicity.*† *Then, an arbitrary projection valued measure P based on $\mathbf{R}^3$ has the property that (V,P) is also a system of imprimitivity for M if and only if there exists a unitary operator A commuting with V such that*

$$P = A \circ P^0 \circ A^{-1}. \tag{217}$$

Proof. If P is defined by (217) for a unitary A lying in the commuting ring of V, it is clear that (V,P) is a system of imprimitivity for M. Conversely, let (V,P) be a system of imprimitivity based on $\mathbf{R}^3$ for M. Then (V,P) and (V,P^0) are equivalent to the systems induced by certain representations s and s^0 of K^*. Since $V \simeq V^s$ and $V \simeq V^{s^0}$, we have $V^s \simeq V^{s^0}$.

---

† Since M is transitive on R^3, P^0 is homogeneous (lemma 9.10). Hence it has a well-defined multiplicity.

From lemma 9.13 applied to the strict cocycles h, $\mathbf{y} \to s(h)$, h, $\mathbf{y} \to s^0(h)$, we conclude that the restrictions of s and s^0 to the diagonal subgroup $T' = \{m_{\zeta,0}\}$ are equivalent. Now a look at the formula (211) shows that the multiplicity of P^0 is the dimension of s^0. Hence, s^0 is finite dimensional. Consequently, s is finite dimensional also. Now, it is clear that two irreducible representations of K^* are equivalent as soon as their restrictions to T' are equivalent. Using the decomposition into irreducible components, one can extend this result easily to *finite dimensional* representations which are not irreducible. This result implies that s and s^0 are equivalent. In other words, the systems of imprimitivity (V,P) and (V,P^0) are themselves equivalent. Consequently there exists a unitary operator A commuting with V which satisfies (217). This finishes the proof of the corollary.

We are now in a position to prove our basic result on the existence of position operators. Let $\mathfrak{S}$ be a system to which is associated the representation U of the inhomogeneous Lorentz group in the Hilbert space $\mathscr{H}$. Then $\mathfrak{S}$ is said to be *localizable* at time $x_0 = 0$ if there exists a projection valued measure P (in $\mathscr{H}$) based on $\mathbf{R}^3$ such that

$$U_{h,a} P_E U_{h,a}^{-1} = P_{\delta(h)[E]+\mathbf{a}}$$

for all $(h, a) \in M$. The following theorem is the main result of this section (cf. Wightman [1]).

Theorem 9.16. *Let the representation U associated with a system $\mathfrak{S}$ be defined by the pair $(X_m{}^+,\pi)$, where $m \geq 0$, and π is a representation of the stability subgroup of L^* at the point $p^{(m)}$, where we write $p^{(m)}$ for $(m,0,0,0)$ when $m > 0$ and $(1,0,0,1)$ when $m = 0$. If $m > 0$, $\mathfrak{S}$ is localizable. If $m = 0$, $\mathfrak{S}$ is localizable if and only if there exists a representation π' of K^* such that π and π' have equivalent restrictions to the subgroup T' of all diagonal matrices of K^*.*

Proof. We note that K^* is the stability subgroup at $p^{(m)}$ for $m > 0$ and E^* is the stability subgroup at $p^{(m)}$ for $m = 0$. Let us now select a strict $(H^*,X_m{}^+)$-cocycle C with values in the Hilbert space $\mathscr{K}$ in which π acts, such that C defines the representation π at $p^{(m)}$. The representation U acts in $\mathscr{H} = \mathscr{L}^2(X_m{}^+,\mathscr{K},\alpha_m{}^+)$ and, for $(h,a) \in M$, we have

$$(U_{h,a}f)(p) = \exp[-i\langle\mathbf{a},\mathbf{p}\rangle]C(h,\delta(h)^{-1}p)f(\delta(h)^{-1}p).$$

We now use the bijection $\mathbf{p} \to ((m^2 + \langle\mathbf{p},\mathbf{p}\rangle)^{1/2},\mathbf{p})$ and reduce this to the form (213). Corollary 9.14 is then applicable and leads to the following necessary and sufficient condition for localizability: $\mathfrak{S}$ is localizable if and only if there exists a representation π' of K^* such that, for almost all $r > 0$, the representation of the stability subgroup at $((m^2+r^2)^{1/2},0,0,r)$, defined by C, has a restriction to the diagonal subgroup T' that is equivalent to the restriction of π' to T'.

Consider now the point $p_{r,m} = ((m^2+r^2)^{1/2},0,0,r)$. Let h_m be the 2×2 diagonal matrix with entries λ and $1/\lambda$, where

$$\lambda = \begin{cases} m^{-1/2}[(m^2+r^2)^{1/2}-r]^{1/2} & (m > 0), \\ r^{-1/2} & (m = 0). \end{cases}$$

It is easy to check that $\delta(h_m)p_{r,m} = p^{(m)}$. If we now write $H^*_{r,m}$ for the stability group at $p_{r,m}$, then $T' = H^*_{r,m} \cap K^*$ and its elements commute with h_m. From the cocycle identities we now find, for $h \in T'$,

$$\begin{aligned} C(h,p^{(m)}) &= C(h_m h h_m{}^{-1},p^{(m)}) \\ &= C(h_m,p_{r,m})C(h,p_{r,m})C(h_m{}^{-1},p^{(m)}) \end{aligned}$$

so that

$$C(h,p_{r,m}) = RC(h,p^{(m)})R^{-1},$$

where

$$R = C(h_m,p_{r,m})^{-1}.$$

In other words, the representations of T' defined by C at the points $((m^2+r^2)^{1/2},0,0,r)$ are all mutually equivalent and are equivalent to the restriction of the representation π to T'. We may now apply the criterion for localizability described in the previous paragraph to conclude that $\mathfrak{S}$ is localizable if and only if for some representation π' of K^*, π, and π' have equivalent restrictions to T'.

For $m>0$, we may take $\pi=\pi'$ always and so any system $\mathfrak{S}$ whose associated representation is defined by the orbit $X_m{}^+$ with $m>0$ is localizable. For $m=0$ the condition obtained is exactly the one stated in the theorem.

Corollary 9.17 (*Newton-Wigner* [1]). *The particles corresponding to the representations* $U^{m,+,j}$ $(m>0, j=0, \frac{1}{2}, 1, \cdots)$ *and* $U^{+,0}$ *are localizable, whereas those corresponding to the representations* $U^{+,\pm n}$ $(n \geq 1)$, *are not. The system corresponding to the representation* $U^{+,1} \oplus U^{+,-1}$ *is localizable. The photon is not localizable.*

Proof. The localizability of the systems corresponding to $U^{m,+,j}$ $(m>0)$ and $U^{+,0}$ is immediate from theorem 9.16. Since K^* has no non-trivial one-dimensional representations, the system corresponding to $U^{+,\pm n}$ is not localizable if $n \geq 1$.

We now come to the systems whose corresponding representations are *reducible*. The representation $U^{+,n} \oplus U^{+,-n}$ is associated with $(X_0{}^+,\pi^{(n)})$, where $\pi^{(n)}$ is the representation

$$m_{\zeta,a} \to \begin{pmatrix} \zeta^n & 0 \\ 0 & \zeta^{-n} \end{pmatrix}$$

of E^* (cf. (84)). For $n=1$, $\pi^{(1)}$ is obviously the restriction of $D^{1/2}$ to T'. Hence the system corresponding to $U^{+,1} \oplus U^{+,-1}$ is localizable. If

$n \geq 2$, the representations D^j whose restrictions to T' contain the character $m_{\zeta,0} \to \zeta^n$ are precisely those for which $2j \geq n$ and $2j - n$ is an integer; the dimension of D^j will then have to be > 2. Hence none of the systems corresponding to $U^{+,n} \oplus U^{+,-n}$, $n \geq 2$, is localizable. For $n = 2$, we obtain the result that the photon is not localizable.

Remark. The representation of K^* of the lowest dimension whose restriction to T' contains the character $m_{\zeta,0} \to \zeta^2$ is D^1, namely the representation $k \to \delta(k)$ $(k \in K^*)$. But the restriction of D^1 to T' contains the characters $m_{\zeta,0} \to \zeta^{\pm 2}$ *and* the trivial character $m_{\zeta,0} \to 1$. In other words, the failure of the photon to be localizable stems, in the last analysis, from the fact that, for a given momentum vector, it has only two independent spin angular momentum states and that, in particular, it does not have any spin angular momentum parallel to its momentum.

The particles with nonzero rest mass have been shown to be localizable, and for these there remains the problem of obtaining the expressions for the position operators. Such formulas were written by Newton and Wigner [1]. We shall now turn to this question. Instead of the position operators which are unbounded it is more convenient to work with the unitary representation of the additive group of $\mathbf{R}^3$ they generate. To be more explicit, let U be the representation associated with a localizable system and let P be a projection valued measure based on $\mathbf{R}^3$ such that the relations (205) are satisfied. We define, for $\mathbf{x} \in \mathbf{R}^3$,

$$B(\mathbf{x}) = \int_{\mathbf{R}^3} \exp[-i\langle \mathbf{x},\mathbf{y}\rangle dP(\mathbf{y})] = \exp[-iq(\mathbf{x})].$$

The $B(\mathbf{x})$ are unitary and $B(\mathbf{x} \to B(\mathbf{x}))$ is a representation of $\mathbf{R}^3$. The $q(\mathbf{x})$ are the position operators. The correspondence (218) between B and P is one-one. It is easily checked that P satisfies (205) if and only if

$$U_{h,a}B(\mathbf{x})U_{h,a}^{-1} = \exp[i\langle \delta(h)\mathbf{x},\mathbf{a}\rangle]B(\delta(h)\mathbf{x}) \tag{218}$$

for all $\mathbf{x} \in \mathbf{R}^3$ and $(h,a) \in M$.

Let us define $W_{h,x}$ and $L(\mathbf{a})$ as follows:

$$W_{h,x} = B(\mathbf{x})U_{h,0},$$

$$L(\mathbf{a}) = U_{1,-a}.$$

Here $(h,a) \in M$. The set of all pairs $(h,\mathbf{x})$ is identified also with M. It follows after a brief calculation that

$$W : h, \mathbf{x} \to W_{h,\mathbf{x}}$$

is a representation of M and that (218) is equivalent to

$$W_{h,\mathbf{x}}L(\mathbf{a})W_{h,\mathbf{x}}^{-1} = \exp[i\langle \delta(h)\mathbf{a},\mathbf{x}\rangle]L(\delta(h)\mathbf{a}). \tag{219}$$

We shall now use (219) to write by inspection the expressions for the $B(\mathbf{x})$. Let $m>0$ and let U be the representation $U^{m,+,0}$ acting in $\mathscr{H}=\mathscr{L}^2(X_m{}^+,\alpha_m{}^+)$; noting that when $a=(0,\mathbf{a})$, $\{a,p\}=-\langle\mathbf{a},\mathbf{p}\rangle$ we find

$$(U_{h,a}f)(p) = \exp[-i\langle\mathbf{a},\mathbf{p}\rangle]f(\delta(h)^{-1}p).$$

The map

$$\eta : p \to \mathbf{p}$$

establishes a homeomorphism of $X_m{}^+$ with $\mathbf{P}^3$. We shall use the notation $\tau_{\mathbf{q}}$ to denote the action in $X_m{}^+$ obtained by lifting the translation in $\mathbf{P}^3$ via η:

$$(220) \qquad \eta(\tau_{\mathbf{q}}(p)) = \mathbf{p}+\mathbf{q}.$$

Since the elements of K^* leave p_0 invariant, an easy calculation shows that

$$\delta(h)\tau_{\mathbf{q}}\delta(h)^{-1} = \tau_{\delta(h)\mathbf{q}} \qquad (h\in K^*).$$

From this it follows that the map

$$(h,\mathbf{q}),p \to \tau_{-\mathbf{q}}(\delta(h)p)$$

converts $X_m{}^+$ into an M-space. In the Hilbert space $\mathscr{H}$, $L(\mathbf{a})$ is the operator of multiplication by the function $\mathbf{p}\to\exp i\langle\mathbf{a},\mathbf{p}\rangle$ and (219) shows that we have a system of imprimitivity for M based on $X_m{}^+$ considered as an M-space. The measure $\alpha_m{}^+$ is, however, not invariant under the "translations" $\tau_{\mathbf{q}}$ and, on noting this, the following expression for the $B(\mathbf{x})$ may be immediately written:

$$(221) \qquad (B(\mathbf{x})f)(p) = \left\{\frac{p_0(p)}{p_0(\tau_{\mathbf{x}}p)}\right\}^{1/2} f(\tau_{\mathbf{x}}p),$$

where p_0 is, as usual, the function $p\to(m^2+\langle\mathbf{p},\mathbf{p}\rangle)^{1/2}$. The reader can verify at once that the $B(\mathbf{x})$ are unitary and that equation (218) is satisfied. The numerical factor on the right side comes as usual from the Radon-Nikodym derivatives. If we write $\mathbf{q}_k$ for the operator $q(\mathbf{e}_k)$, where $\mathbf{e}_k$ is the vector, all of whose components except the kth one vanish and whose kth component is unity, we obtain from (221) the equation

$$(222) \qquad i\mathbf{q}_k f = -\frac{\partial}{\partial p_k}f+\frac{p_k}{2p_0{}^2}f;$$

here p_1, p_2, p_3 are global coordinates on $X_m{}^+$ and the $\partial/\partial p_k$ are the corresponding derivatives, while f is a smooth function vanishing outside a compact set.

The most general B satisfying (218) can now be determined by an application of corollary 9.15. Let W be any unitary operator commuting with the representation U restricted to M. Commutativity with all

$U_{1,a}$ ($a_0=0$) leads at once to the fact that W is multiplication by a Borel function, say u, such that $|u(p)|=1$ for all p. The commutativity with all $U_{h,0}$ ($h \in K^*$) now forces u to be K^*-invariant and hence to be a function of p_0. Hence the most general B such that $\mathbf{x} \to B(\mathbf{x})$ is a unitary representation of $\mathbf{R}^3$ satisfying (218) is of the form

$$(B(\mathbf{x})f)(p) = \exp[i(k(p_0)(p) - k(p_0)(\tau_{\mathbf{x}}p))]\left\{\frac{p_0(p)}{p_0(\tau_{\mathbf{x}}p)}\right\}^{1/2} f(\tau_{\mathbf{x}}p),$$

where k is an arbitrary real valued Borel function on the positive reals.

We shall now proceed to the case of nonzero spin. Let us take $U = U^{m,+,j}$ in the form (154), acting in the Hilbert space $\mathscr{H}$ of sections of the bundle $B_m^{+,j}$. Let us assume first that $j=\frac{1}{2}$. It is natural to assume for $B(\mathbf{x})$ an expression of the form

$$(B(\mathbf{x})f)(p) = \left\{\frac{p_0(p)}{p_0(\tau_{\mathbf{x}}p)}\right\}^{1/2} S(R(\mathbf{x},p))f(\tau_{\mathbf{x}}p),$$

where $R(\mathbf{x},p) \in H^*$ and $S(h \to S(h))$ is the usual spin representation of H^*. In order that $S(R(\mathbf{x},p))$ may transform the fiber at $\tau_{\mathbf{x}}p$ to the fiber at p, it is enough to assume that $\delta(R(\mathbf{x},p)^*)$ maps p onto $\tau_{\mathbf{x}}p$. In order that B may be a representation satisfying (218), we should require that the map $\mathbf{x}, p \to R(\mathbf{x},p)$ satisfy the conditions:

$$(223)\quad \begin{aligned} &\text{(i)} && hR(\mathbf{x},p)h^{-1} = R(\delta(h)\mathbf{x},\delta(h)p) \qquad (h \in K^*),\\ &\text{(ii)} && R(\mathbf{x}+\mathbf{x}',p) = R(\mathbf{x},p)R(\mathbf{x}',\tau_{\mathbf{x}}p),\\ &&& R(\mathbf{0},p) = 1,\\ &\text{(iii)} && \delta(R(\mathbf{x},p)^*)p = \tau_{\mathbf{x}}p. \end{aligned}$$

An easy calculation using (152) shows that each $B(\mathbf{x})$ will be unitary.

We shall now "solve" equations (223) for the $R(\mathbf{x},p)$. $\mathbf{R}^3$ acts freely and transitively on $X_m{}^+$ by the action $\mathbf{x}, p \to \tau_{\mathbf{x}}p$, and (ii) of (223) just asserts that $\mathbf{x}, p \to R(\mathbf{x},p)^{-1}$ is a strict $(\mathbf{R}^3, X_m{}^+)$-cocycle. Therefore, by lemma 5.23 we may write

$$R(\mathbf{x},p) = L(p)^{-1}L(\tau_{\mathbf{x}}p),$$

where $p \to L(p)$ is an arbitrary function. The condition (i) then becomes, on simplification,

$$L(\delta(h)p)hL(p)^{-1} = L(\delta(h)\tau_{\mathbf{x}}p)hL(\tau_{\mathbf{x}}p)^{-1}.$$

This shows that the function $p \to L(\delta(h)p)hL(p)^{-1}$ is $\tau_{\mathbf{x}}$-invariant for all $\mathbf{x}$, and therefore must be a constant, possibly depending on h:

$$L(\delta(h)p)hL(p)^{-1} = C_h \qquad (h \in K^*).$$

Equation (iii) of (223) can be analyzed similarly. We have, on substituting the expression for R in (iii),

$$\delta(L(p)^{*-1})p = \delta(L(\tau_{\mathbf{x}}p)^{*-1})(\tau_{\mathbf{x}}p).$$

This shows that $p \to \delta(L(p)^{*-1})p$ must be a constant also. If we use the map

$$p \to A(p) = p_0\sigma_0 + p_1\sigma_1 + p_2\sigma_2 + p_3\sigma_3$$

which sends elements of $\mathbf{P}^4$ into 2×2 Hermitian matrices, and write the condition that $p \to \delta(L(p^*)^{-1})p$ has the same value for all p, we get

$$L(p)^{*-1}A(p)L(p)^{-1} = mL(p^{(m)})^{*-1}L(p^{(m)})^{-1},$$

where $p^{(m)} = (m,0,0,0)$. (iii) of (223) therefore can be replaced by

$$A(p) = m\{L(p^{(m)})^{-1}L(p)\}^*\{L(p^{(m)})^{-1}L(p)\}.$$

It is now easy to determine a solution L. We define $L(p^{(m)}) = 1$ and

$$L(p) = m^{-1/2}A(p)^{1/2}. \tag{224}$$

Note that for $p \in X_m{}^+$ $(m > 0)$, $A(p)$ is positive definite so that $L(p)$ is well defined and $\det L(p) = 1$. We then obtain, for R, the following expression:

$$R(\mathbf{x},p) = A(p)^{-1/2}A(\tau_{\mathbf{x}}p)^{1/2}. \tag{225}$$

We note that (i) of (223) is automatically satisfied $(C_h = h)$. Collecting all this, we obtain

$$(B(\mathbf{x})f)(p) = \left\{\frac{p_0(p)}{p_0(\tau_{\mathbf{x}}p)}\right\}^{1/2} S(A(p)^{-1/2}A(\tau_{\mathbf{x}}p)^{1/2})f(\tau_{\mathbf{x}}p). \tag{226}$$

Actually we now use (226) to *define* B. It is not difficult to check that the $B(\mathbf{x})$ are unitary and that $\mathbf{x} \to B(\mathbf{x})$ is a unitary representation of $\mathbf{R}^3$ satisfying (218). For higher spin, we replace (226) by

$$(B(\mathbf{x})f)(p) = \left\{\frac{p_0(p)}{p_0(\tau_{\mathbf{x}}p)}\right\}^{1/2} S_N(A(p)^{-1/2}A(\tau_{\mathbf{x}}p)^{1/2})f(\tau_{\mathbf{x}}p), \tag{227}$$

where $S_N = S \otimes \cdots \otimes S$ $(N = 2j$ factors).

We can now write the infinitesimal form of (226). To do this it is necessary, among other things, to compute

$$\mathfrak{R}(\mathbf{x},p) = \left\{\frac{d}{dt} S(A(p)^{-1/2}A(\tau_{t\mathbf{x}}p)^{1/2})\right\}_{t=0}.$$

Obviously, if we write

$$\mathfrak{N}(\mathbf{x},p) = A(p)^{-1/2}\left\{\frac{d}{dt} A(\tau_{t\mathbf{x}}p)^{1/2}\right\}_{t=0},$$

then $\mathfrak{N}(\mathbf{x},p)$ is a matrix of zero trace and

$$\mathfrak{R}(\mathbf{x},p) = \dot{S}(\mathfrak{N}(\mathbf{x},p)),$$

where $\dot{S}$ is the differential of the representation S. From (i) of (223), we have:

$$S(h)\mathfrak{R}(\mathbf{x},p)S(h)^{-1} = \mathfrak{R}(\delta(h)\mathbf{x},\delta(h)p) \qquad (h \in K^*), \tag{228}$$

so that $\mathfrak{R}$ is uniquely determined by its values for all $\mathbf{x}$ and all p of the form $p = (p_0,0,0,p_3)$. For fixed $\mathbf{x}$ and such p, we find, on writing

$$u = \frac{d}{dt}\{A(\tau_{t\mathbf{x}}p)^{1/2}\}_{t=0}$$

and differentiating the equation $[A(\tau_{t\mathbf{x}}p)^{1/2}]^2 = A(\tau_{t\mathbf{x}}p)$, that

$$A(p)^{1/2}u + uA(p)^{1/2} = \begin{pmatrix} x_3(p_0+p_3)/p_0 & x_1 - ix_2 \\ x_1 + ix_2 & -x_3(p_0-p_3)/p_0 \end{pmatrix}.$$

Using the fact that u is Hermitian we can compute it from this equation. Then $\mathfrak{R}(\mathbf{x},p)$ reduces, after some simplification, to

$$\mathfrak{R}(\mathbf{x},p) = \begin{pmatrix} x_3/2p_0 & (x_1 - ix_2)/p_0 + p_3 + m \\ (x_1 + ix_2)/p_0 - p_3 + m & -x_3/2p_0 \end{pmatrix}.$$

Expressing $\mathfrak{R}(\mathbf{x},p)$ as a real linear combination of the σ_r and the $i\sigma_r$, and using (54) and (134), we find, after some calculation, that

$$\mathfrak{R}(\mathbf{x},p) = \frac{1}{2m}\sum_{r=1}^{3} x_r\gamma_0\gamma_r - \frac{x_3p_3}{2mp_0(p_0+m)}\,p_3\gamma_0\gamma_3 - \frac{1}{2m(p_0+m)}\{x_2p_3\gamma_2\gamma_3 - x_1p_3\gamma_3\gamma_1\}.$$

This formula is valid for $p = (p_0,0,0,p_3)$. A look at the requirement (228) of covariance now enables us to write the expression for $\mathfrak{R}(\mathbf{x},p)$ in the general case:

$$\mathfrak{R}(\mathbf{x},p) = \frac{-\langle \mathbf{x},\mathbf{p}\rangle}{2mp_0(p_0+m)}\sum_{j=1}^{3} p_j\gamma_0\gamma_j + \frac{1}{2m}\sum_{j=1}^{3} x_j\gamma_0\gamma_j - \frac{1}{2m(p_0+m)}\sum_{1\le j<k\le 3}(x_jp_k - x_kp_j)\gamma_j\gamma_k. \tag{229}$$

The fact that $\mathfrak{R}(\mathbf{x},p)$ as defined above satisfies (228) for all unitary h is straightforward. It is also trivial that, for $p_1 = p_2 = 0$ (229) reduces to the earlier expression. We omit the verifications of these. We finally obtain, from (229), the following expressions for the position operators $q(\mathbf{x})$:

$$iq(\mathbf{x})f = -\sum_{k=1}^{3} x_k\frac{\partial}{\partial p_k}f + \frac{\langle \mathbf{x},\cdot\rangle}{2p_0{}^2}f - \mathfrak{R}(\mathbf{x},\cdot)f \tag{230}$$

for any smooth section f vanishing outside a compact set. The generalization of (230) to arbitrary j is trivial. If we define $\mathfrak{R}^{\nu}(\mathbf{x},p)$ by (229) with γ_j^{ν} replacing γ_j (cf. (147), for the meaning of γ_j^{ν}), and write

$$\mathfrak{R}(\mathbf{x},p) = \sum_{\nu=1}^{N} \mathfrak{R}^{\nu}(\mathbf{x},p),$$

then (230) once again gives the expressions for the $iq(\mathbf{x})$ when $j=\frac{1}{2}N$ is arbitrary.

We shall now take up the question of uniqueness. We consider first $j=\frac{1}{2}$. In view of corollary 9.15, we may replace the unitary representation B by $W \circ B \circ W^{-1}$, where W is a unitary operator which commutes with all $U_{h,a}$ $(h,a) \in M$, without destroying the imprimitivity equations (205). Now, for any such W, the fact that the $U_{1,a}$ $(a_0=0)$ commute with it shows that there exists a Borel map

$$w : p \rightarrow w(p)$$

of $X_m{}^+$ into the unitary group of $\mathbf{C}^4$ such that, for each p, $w(p)$ leaves the fiber of $B_m^{+,1,2}$ at p invariant, and $(W\varphi)(p) = w(p)\varphi(p)$ for almost all p and all φ of the Hilbert space. The commutativity with all $U_{h,0}$ $(h \in K^*)$ now leads to the equations

$$w(\delta(h)p)S(h) = S(h)w(p)$$

for each h for almost all p. For $p=(p_0,0,0,p_3)$ with $p_3>0$, the $S(h)$ for those h with $\delta(h)p=p$ are diagonal. It is then not too difficult to determine the w which satisfy these equations. We choose a Borel function v, $r \rightarrow v(r)$, of the positive reals into the group of 2×2 diagonal matrices of K^* arbitrarily and define w by

$$w(p_r) = \begin{pmatrix} v(r) & 0 \\ 0 & v(r) \end{pmatrix} \qquad (p_r = ((m^2+r^2)^{1/2},0,0,r),\ v(0) = 1),$$

$$w(\delta(h)p_r) = S(h)w(p_r)S(h)^{-1} \qquad (h \in K^*).$$

It is clear that w is well defined and meets the requirements. For a given w, the unitary operator W corresponding to it is given by

$$(W\varphi)(p) = w(p)\varphi(p) \qquad (p \in X_m{}^+).$$

The unitaries thus obtained exhaust the commuting ring in question. The case of higher spin is handled similarly.

It is noteworthy that the position observables cannot be uniquely determined. Formulas (222) and (230), however, seem to give the simplest of the possible choices. For the choices represented by (222) and (230), we can calculate the *velocity* observables by differentiating the position observables with respect to time. Since (cf. Chapter XI) these are their commutators with energy, we find, using (230) and writing $q_k = q(\mathbf{e}_k)$,

$$v_k = \dot{q}_k = p_k/p_0. \tag{231}$$

The velocities are conserved in time. Formulas (231) are the same as those in classical relativistic mechanics.

8. GALILEAN RELATIVITY

Our concern so far has been exclusively with the relativity prescribed by the Lorentz group. Clearly the whole circle of problems can be analyzed under the over-all assumption that the space-time group is the inhomogeneous group of *Galilean* transformations. We shall indicate briefly in this section the modifications that are to be made in the foregoing analysis when we change the underlying group. To describe the free particles under the assumption of Galilean relativity, we must describe and classify up to physical equivalence the irreducible (possibly projective) representations of the inhomogeneous Galilean group. We carry out our discussion along the same lines as we did for the Lorentz case.

Let us now consider the universal covering group of the inhomogeneous Galilean group (cf. (112) through (123) of Chapter VII). This consists of the quadruples

$$(h,\eta,\mathbf{v},\mathbf{u}) \qquad (h \in K^*, \eta \in \mathbf{R}^1, \mathbf{v}, \mathbf{u} \in \mathbf{R}^3),$$

with the multiplication law

$$(h,\eta,\mathbf{v},\mathbf{u})(h',\eta',\mathbf{v}',\mathbf{u}') = (hh', \eta+\eta', \delta(h)\mathbf{v}'+\mathbf{v}, \delta(h)\mathbf{u}'+\mathbf{u}+\eta'\mathbf{v}),$$

where h, $h' \in K^*$, η, $\eta' \in \mathbf{R}^1$, $\mathbf{u}$, $\mathbf{v} \in \mathbf{R}^3$. The space-time translations $(1,\eta,\mathbf{0},\mathbf{u})$ form an abelian normal subgroup; the group of the homogeneous Galilean transformations $(h,0,\mathbf{v},\mathbf{0})$ acts on the translations (via inner automorphisms) as follows:

$$(h,0,\mathbf{v},\mathbf{0}) : (1,\eta,\mathbf{0},\mathbf{u}) \rightarrow (1,\eta,\mathbf{0},\delta(h)\mathbf{u}+\eta\mathbf{v}).$$

This gives, of course, the semidirect product structure of the inhomogeneous Galilean group. We introduce "momentum space" $\mathbf{P}^4$ of points $(p_0,\mathbf{p})$ as the dual of the vector space of space-time translations $(1,\eta,\mathbf{0},\mathbf{u})$. Each $(p_0,\mathbf{p})$ gives rise to the character

$$(1,\eta,\mathbf{0},\mathbf{u}) \rightarrow \exp[i(p_0\eta+\langle\mathbf{p},\mathbf{u}\rangle)],$$

where $\langle . , . \rangle$ is the Euclidean inner product in $\mathbf{R}^3$. The adjoint action of $(h,0,\mathbf{v},\mathbf{0})$ in $\mathbf{P}^4$ can now be computed; an easy calculation shows this to be:

$$(h,0,\mathbf{v},\mathbf{0}) : (p_0,\mathbf{p}) \rightarrow (p_0-\langle\mathbf{v},\delta(h)\mathbf{p}\rangle, \delta(h)\mathbf{p}). \tag{232}$$

From (232) one can easily obtain the orbits for the homogeneous group in $\mathbf{P}^4$. We claim that the orbits are

(233) $$Z_{p_0} = \{(p_0,\mathbf{0})\} \qquad (-\infty < p_0 < \infty),$$

and

(234) $$X_r = \{(p_0,\mathbf{p}) : \langle \mathbf{p},\mathbf{p}\rangle = r^2\} \qquad (r > 0).$$

That Z_{p_0}, which consists of a single point, is an orbit is trivial. On the other hand, (232) shows that X_r is invariant. We shall show that any point of X_r is obtained from $(0,(0,0,r))$ by applying a suitable element $(h,0,\mathbf{v},\mathbf{0})$. In fact, (232) shows that any point $(0,\mathbf{p})$ with $\langle \mathbf{p},\mathbf{p}\rangle = r^2$ can be so obtained; to pass from $(0,\mathbf{p})$ to $(p_0,\mathbf{p})$ we choose $\mathbf{v} \in R^3$ so that $\langle \mathbf{v},\mathbf{p}\rangle = -p_0$ and use $(1,0,\mathbf{v},\mathbf{0})$. Note that X_r "splits" into $R^1 \times S^r$, where S^r is the sphere of radius r in $\mathbf{R}^3$. An easy calculation based on Fubini theorem shows that on X_r, the measure $dp_0 d\sigma^r$ is invariant, dp_0 being Lebesgue measure on $\mathbf{R}^1$ and σ^r the unique normalized rotation invariant measure on S^r. The stability group at $(p_0,\mathbf{0})$ is the entire homogeneous group. The stability group at $(0,(0,0,r))$ consists of all elements $(h,0,\mathbf{v},\mathbf{0})$, where

$$h = m_{\zeta,0} = \begin{pmatrix} \zeta & 0 \\ 0 & \zeta^{-1} \end{pmatrix} \qquad (|\zeta| = 1)$$

and

$$\mathbf{v} = (v_1,v_2,0) \qquad (v_1,\ v_2 \text{ arbitrary}).$$

The product of $(m_{\zeta,0},0,\mathbf{v},\mathbf{0})$ and $(m_{\zeta',0},0,\mathbf{v}',\mathbf{0})$ is $(m_{\zeta\zeta',0},0,\mathbf{v}'',\mathbf{0})$, where

$$\mathbf{v}'' = \delta(m_{\zeta,0})\mathbf{v}' + \mathbf{v}.$$

On writing

$$z(\mathbf{v}) = v_1 - iv_2,$$

we obtain for $z(\mathbf{v}'')$:

$$z(\mathbf{v}'') = \zeta^2 z(\mathbf{v}') + z(\mathbf{v}).$$

In other words, the stability subgroup at $(0,(0,0,r))$ is isomorphic to the semidirect product of $\mathbf{C}$ and T under the mapping $(m_{\zeta,0},0,\mathbf{v},\mathbf{0}) \to (\zeta,z(\mathbf{v}))$, wherein the product in $\mathbf{T} \times \mathbf{C}$ is determined according to

(235) $$(\zeta,z)(\zeta',z') = (\zeta\zeta',\zeta^2 z' + z).$$

The orbit structure of $\mathbf{P}^4$ is obviously smooth. For the stability subgroup at $(0,(0,0,r))$ we have already described the representations. These are the π_n,

$$\pi_n : (m_{\zeta,0},0,\mathbf{v},\mathbf{0}) \to \zeta^n \qquad (n = 0,\ \pm 1,\ \pm 2,\cdots)$$

and the π_ρ and π_ρ' $(\rho > 0)$; the latter act in $\mathscr{L}^2(T)$ and are described by (cf. (98) and (99))

$$(\pi_\rho(m_{\zeta,0},0,\mathbf{v},\mathbf{0})f)(t) = \exp[i\rho \operatorname{Re}(t\zeta^{-1}z(\mathbf{v})^*)]f(\zeta^{-2}t),$$
$$(\pi_\rho'(m_{\zeta,0},0,\mathbf{v},\mathbf{0})f)(t) = \zeta \exp[i\rho \operatorname{Re}(t\zeta^{-1}z(\mathbf{v})^*)]f(\zeta^{-2}t).$$

The representations L^{π} associated with any one of these irreducibles, say π, may now be written explicitly. We only note that

$$(L^{\pi}_{(1,\eta,0,\mathbf{u})}f)(p_0,\mathbf{p}) = \exp[i(\eta p_0+\langle\mathbf{u},\mathbf{p}\rangle)]f(p_0,\mathbf{p}), \tag{236}$$

f belonging to a Hilbert space of scalar or vector valued functions on X_r. The representations associated with Z_{p_0} are

$$(h,\eta,\mathbf{v},\mathbf{u}) \to \exp(ip_0\eta)\pi(h,0,\mathbf{v},\mathbf{0}), \tag{237}$$

where π is an arbitrary irreducible representation of the homogeneous group.

It is remarkable that none of these representations can represent any physical particle. We shall see now why this is so. Formula (237) shows that the time translations act as *scalars* and hence leave every ray fixed—there is no temporal evolution. Hence it becomes necessary to give up these representations. We come next to the representations L^{π} associated with the orbits X_r $(r>0)$. Formula (236) shows that the momentum operators B_1, B_2, B_3 corresponding to the translations along the axes in space are the operators of multiplication by p_1, p_2, and p_3. But $p_1{}^2+p_2{}^2+p_3{}^2=r^2$ and hence

$$B_1{}^2+B_2{}^2+B_3{}^2 = r^2 1, \tag{238}$$

which gives an unphysical relation between the three momenta. Moreover, it can be quickly seen that the systems corresponding to these representations are not localizable either. To see this, let us write V for the restriction of any one of these representations L^{π} to the subgroup of translations in space only and let us assume that there exists a projection valued measure $P(E \to P_E)$ based on the σ-algebra of Borel sets of $\mathbf{R}^3$ acting in the Hilbert space $\mathscr{H}$ of V such that

$$V_{\mathbf{u}}P_E V_{\mathbf{u}}{}^{-1} = P_{E+\mathbf{u}}$$

for all $\mathbf{u}\in\mathbf{R}^3$ and $E\subseteq\mathbf{R}^3$. From theorem 6.17, we know that there exists a unitary isomorphism W of $\mathscr{H}$ with $\mathscr{L}^2(\mathbf{R}^3,\mathscr{K},d\mathbf{x})$, where $\mathscr{K}$ is a Hilbert space, such that

$$((WV_{\mathbf{u}}W^{-1})f)(\mathbf{x}) = f(\mathbf{x}-\mathbf{u})$$

for all $f\in\mathscr{L}^2(\mathbf{R}^3,\mathscr{K},d\mathbf{x})$. Now, $\mathbf{u}\to V_{\mathbf{u}}$ is a representation of the abelian group $\mathbf{R}^3$ and hence there exists a unique projection valued measure Q on the space of vectors $\mathbf{p}$ such that

$$V_{\mathbf{u}} = \int_{\mathbf{P}^3} \exp[i\langle\mathbf{u},\mathbf{p}\rangle]dQ(\mathbf{p}).$$

From the uniqueness of Q and from formula (236) we see that Q vanishes for the complement of the set $\{\mathbf{p} : \langle \mathbf{p},\mathbf{p}\rangle = r^2\}$; in fact, (236) shows that in the Hilbert space

$$\mathscr{H}' = \mathscr{L}^2(X_r,\mathscr{K},d\sigma \times dp_0),$$

$\mathscr{K}$ being the space of π,. $V_{\mathbf{u}}$ acts as multiplication by the function $\mathbf{p} \to \exp i\langle \mathbf{u},\mathbf{p}\rangle$ and hence Q_E must be the operator of multiplication by χ_E, proving that Q vanishes for the complement of the set $\langle \mathbf{p},\mathbf{p}\rangle = r^2$. Therefore, the projection valued measure WQW^{-1} also vanishes for the complement of the set $\{\mathbf{p} : \langle \mathbf{p},\mathbf{p}\rangle = r^2\}$. On the other hand, we use the Fourier-Plancherel isomorphism (cf. (209))

$$\mathscr{F} : f \to \hat{f}$$

of $\mathscr{L}^2(\mathbf{R}^3,\mathscr{K},d\mathbf{x})$ with $\mathscr{L}^2(\mathbf{P}^3,\mathscr{K},d\mathbf{p})$ to find

$$(\mathscr{F}WV_{\mathbf{u}}W^{-1}\mathscr{F}^{-1}\hat{f})(\mathbf{p}) = \exp[-i\langle \mathbf{u},\mathbf{p}\rangle]\hat{f}(\mathbf{p}).$$

This implies that the measure class of the projection valued measure WQW^{-1} is that of *Lebesgue measure*, which is a contradiction, as we have already shown that this measure class is supported by the set $\{\mathbf{p} : \langle \mathbf{p},\mathbf{p}\rangle = r^2\}$ of zero Lebesgue measure. Therefore, the systems corresponding to the representations L^π are not localizable.

Our analysis therefore shows that we must consider the *projective* representations of the inhomogeneous Galilean group in order to be able to build a theory of Galilean invariant particles. We shall now indicate the details of this theory. We shall use the results and the theory of multipliers as developed in Chapter VII. According to the results we have obtained there, any nonexact multiplier of the covering group of the Galilean group is similar to the multiplier m_τ which is given by

$$(239) \qquad m_\tau : r, r' \to \exp i\,\frac{\tau}{2}\,[\langle \mathbf{u},\delta(h)\mathbf{v}'\rangle - \langle \mathbf{v},\delta(h)\mathbf{u}'\rangle + \eta'\langle \mathbf{v},\delta(h)\mathbf{v}'\rangle],$$

where

$$r = (h,\eta,\mathbf{v},\mathbf{u}), \qquad r' = (h',\eta',\mathbf{v}',\mathbf{u}'),$$

and

$$\tau \neq 0$$

is a real number. Given m_τ, we can construct the associated extension of the inhomogeneous group in the usual way. The elements of this extension are pairs $(r;\, \zeta)$ with $|\zeta| = 1$, and the group operation is given by

$$(r;\, \zeta)(r';\, \zeta') = (rr';\, \zeta\zeta' m_\tau(r,r')).$$

We denote this group by G^τ. Now $m_\tau(r,r') = 1$ when $\mathbf{v} = \mathbf{v}' = 0$ and so

$$(1,\eta,\mathbf{0},\mathbf{u};\, \zeta)(1,\eta',\mathbf{0},\mathbf{u}';\, \zeta') = (1,\eta+\eta',\mathbf{0},\mathbf{u}'+\mathbf{u};\, \zeta\zeta'),$$

so that the set of all elements of the form $(1,\eta,\mathbf{0},\mathbf{u};\zeta)$ forms an abelian group. We denote this group by A^τ and, when convenient, its typical element by $(\eta,\mathbf{u};\zeta)$. Further, $m_\tau(r,r')=1$ when $\mathbf{u}=\mathbf{u}'=0$ and $\eta'=0$ so that

$$(h,0,\mathbf{v},\mathbf{0};1)(h',0,\mathbf{v}',\mathbf{0};1) = (hh',0,\delta(h)\mathbf{v}'+\mathbf{v},0;1),$$

showing that the set of all elements $(h,0,\mathbf{v},\mathbf{0};1)$ forms a subgroup of G^τ. We write H^τ for this subgroup and, when convenient, denote its typical element by $(h,\mathbf{v};1)$. We have:

$$\left(\eta,\mathbf{u};\zeta\exp\left[-i\frac{\tau}{2}\langle\mathbf{u},\mathbf{v}\rangle\right]\right)(h,\mathbf{v};1) = (h,\eta,\mathbf{v},\mathbf{u};\zeta). \tag{240}$$

G^τ is thus a semidirect product of A^τ and H^τ. The action of H^τ on A^τ via inner automorphisms is given by

$$(h,\mathbf{v};1):(\eta,\mathbf{u};\zeta)\rightarrow(\eta,\delta(h)\mathbf{u}+\eta\mathbf{v},\zeta'), \tag{241}$$

where

$$\zeta' = \zeta\exp -i\frac{\tau}{2}[2\langle\mathbf{v},\delta(h)\mathbf{u}\rangle+\eta\langle\mathbf{v},\mathbf{v}\rangle]. \tag{242}$$

We are interested in describing those irreducible representations of G^τ which map the elements $(1,0,0,0;\zeta)$ into ζ^{-1} times the identity, for these are precisely the representations which give rise to the m_τ-representations of the inhomogeneous Galilean group (theorem 7.16). Fortunately, G^τ is itself a semidirect product of A^τ and H^τ, and so we can determine its irreducible representations using the theory of Section 8, Chapter VI. To this end, we must determine the dual group $\widehat{A^\tau}$ of A^τ and classify the H^τ orbits in $\widehat{A^\tau}$. A typical element of $\widehat{A^\tau}$ will be written as $(p_0,\mathbf{p};n)$ where $p_0\in R^1$, $\mathbf{p}\in R^3$, and n is an integer; we shall associate with it the character

$$(\eta,\mathbf{u};\zeta)\rightarrow\zeta^n\exp[i(\eta p_0+\langle\mathbf{u},\mathbf{p}\rangle)].$$

A brief calculation based on (241) shows that the adjoint action of the element $(h,\mathbf{v};1)$ in $\widehat{A^\tau}$ is given by

$$(h,\mathbf{v};1):(p_0,\mathbf{p},n)\rightarrow(p_0',\mathbf{p}',n), \tag{243}$$

where

$$\begin{aligned} p_0' &= p_0-\langle\mathbf{v},\delta(h)\mathbf{p}\rangle-\tfrac{1}{2}n\tau\langle\mathbf{v},\mathbf{v}\rangle, \\ \mathbf{p}' &= \delta(h)\mathbf{p}+n\tau\mathbf{v}. \end{aligned} \tag{244}$$

The reader may observe the difference between this action and the action (232).

It is quite simple to determine the orbits and stability groups associated with the action (244). A brief calculation shows that

$$(245) \qquad \langle \mathbf{p},\mathbf{p}\rangle + 2n\tau p_0 = \langle \mathbf{p}',\mathbf{p}'\rangle + 2n\tau p_0'.$$

In other words, the function $(p_0,\mathbf{p};\, n) \rightarrow \langle \mathbf{p},\mathbf{p}\rangle + 2n\tau p_0$ is invariant under H^τ for each fixed n. Let us now define the sets $Z_{n,\rho}$ by

$$(246) \qquad Z_{n,\rho} = \{(p_0,\mathbf{p};\, n) : \langle \mathbf{p},\mathbf{p}\rangle + 2n\tau p_0 = \rho\};$$

here n is an integer, ρ a real number, n and ρ are arbitrary except when $n=0$, for which case $\rho \geq 0$. The $Z_{n,\rho}$ are invariant under H^τ and

$$\widehat{A^\tau} = \bigcup_{n,\rho} Z_{n,\rho}.$$

When $n=0$, (244) reduces to (232) and we find that the sets $Z_{0,\rho}$ $(\rho>0)$ are orbits. The set $Z_{0,0}$ is not an orbit; $Z_{0,0}=\{(p_0,\mathbf{0};\ 0) : p_0 \in \mathbf{R}^1\}$, and each point of it is invariant under H^τ. For $n \neq 0$, and arbitrary ρ, $Z_{n,\rho}$ is an orbit. In fact, the point

$$(247) \qquad a_{n,\rho} = ((2n\tau)^{-1}\rho,\mathbf{0};\, n)$$

lies in $Z_{n,\rho}$, and one computes quickly from (243) that the element $(1,(n\tau)^{-1}\mathbf{p};\, 1)$ of H^τ sends $a_{n,\rho}$ to the point $((2n\tau)^{-1}(\rho - \langle \mathbf{p},\mathbf{p}\rangle),\mathbf{p};\, n)$ which is, of course, the general point of $Z_{n,\rho}$. The stability subgroup of H^τ at $a_{n,\rho}$ is the group of all elements of the form

$$(h,\mathbf{0};\, 1)$$

which is isomorphic to K^* (the unitary group). The map

$$\xi : \mathbf{p} \rightarrow ((2n\tau)^{-1}(\rho - \langle \mathbf{p},\mathbf{p}\rangle),\mathbf{p};\, n)$$

is a homeomorphism of $Z_{n,\rho}$ $(n \neq 0)$ with $\mathbf{P}^3$ and allows us to treat p_1, p_2, and p_3 as coordinates on $Z_{n,\rho}$. The element $(h,\mathbf{0};\, 1)$ sends $(p_0,\mathbf{p};\, n)$ to $(p_0,\delta(h)\mathbf{p};\, n)$ while the element $(1,\mathbf{v};\, 1)$ sends $(p_0,\mathbf{p};\, n)$ to the element $(p_0 - \langle \mathbf{v},\mathbf{p}\rangle - \frac{1}{2}n\tau\langle \mathbf{v},\mathbf{v}\rangle,\ \mathbf{p}+n\tau\mathbf{v};\, n)$. From this it follows easily through an application of Fubini theorem that the measure $d\mathbf{p}$ is invariant under H^τ. Finally it is obvious that the entire orbit structure of $\widehat{A^\tau}$ is smooth.

Unless $n=-1$, the orbit $Z_{n,\rho}$ will not lead to an m_τ-representation of G. If Z is an orbit and U an irreducible representation of G^τ associated with Z, then U acts on a Hilbert space $\mathscr{H}$ of functions defined on Z with values in some Hilbert space $\mathscr{K}$. From the general theory of Chapter VI, we know that for any $a \in A^\tau$, U_a will be the operator of multiplication by the function $\hat{z} \rightarrow \hat{z}(a)$ $(\hat{z} \in Z)$. In particular, if $Z \subseteq Z_{n,\rho}$, the element $(0,\mathbf{0};\, \zeta)$ of A^τ will be mapped, in any irreducible representation associated with Z, into the operator $\zeta^n I$ (I being the identity operator). Consequently, the

m_τ-representations of G are precisely those that are associated with the orbits $Z_{-1,\rho}$, ρ arbitrary.

We shall now describe these representations. Exactly as in the case of the Euclidean group (cf. Chapter VIII), we find that, if

$$s : h \rightarrow s(h)$$

is an irreducible representation of the unitary group K^* in a (finite dimensional) Hilbert space $\mathscr{K}$, the map $(h,\mathbf{v}; 1), \hat{z} \rightarrow s(h)$ $(\hat{z} \in Z_{-1,\rho})$ is a strict cocycle which defines the representation $(h,\mathbf{0}; 1) \rightarrow s(h)$ of the stability group at $a_{-1,\rho}$. If we denote by $\mathscr{H}$ the Hilbert space

$$\mathscr{H} = \mathscr{L}^2(\mathbf{P}^3,\mathscr{K},d\mathbf{p}), \tag{248}$$

then the representation $U^{\tau,s,\rho}$ of G^τ, associated with $Z_{-1,\rho}$ and s, is given by

$$(U^{\tau,s,\rho}_{r,\zeta}f)(\mathbf{p}) = \Omega' s(h) f(\delta(h)^{-1}(\mathbf{p}+\tau\mathbf{v})), \tag{249}$$

where $r=(h,\eta,\mathbf{v},u)$ and

$$\Omega' = (\zeta e^{(i\rho\eta)/2\tau})^{-1} \exp i\left[\left\langle \mathbf{u}, \mathbf{p}+\frac{\tau}{2}\mathbf{v}\right\rangle + \frac{\eta}{2\tau}\langle \mathbf{p},\mathbf{p}\rangle\right]; \tag{250}$$

for this we have only to use the map ξ identifying $Z_{-1,\rho}$ with $\mathbf{P}^3$, the equation (240) which displays explicitly the semidirect product structure of G^τ, the equation (244) which gives the adjoint action of H^τ, and substitute in the general formula (132) of Chapter VI.

The map $r \rightarrow U^{\tau,s,\rho}_{r,1}$ will then be the irreducible m_τ representation of the inhomogeneous Galilean group associated with $U^{\tau,s,\rho}$. We now observe that the number ρ enters (250) only in the factor $e^{-i\rho\eta/2\tau}$ and hence may be dropped when we describe the corresponding *projective* representation of the inhomogeneous Galilean group. We shall write $V^{\tau,s}$ for this irreducible, projective representation. It acts on $\mathscr{H}$ and

$$(V^{\tau,s}_r f)(\mathbf{p}) = \Omega s(h) f(\delta(h)^{-1}(\mathbf{p}+\tau\mathbf{v})) \tag{251}$$

with

$$\Omega = \exp i\left[\left\langle \mathbf{u}, \mathbf{p}+\frac{\tau}{2}\mathbf{v}\right\rangle + \frac{\eta}{2\tau}\langle \mathbf{p},\mathbf{p}\rangle\right]. \tag{252}$$

In this form, the multiplier for $V^{\tau,s}$ will be similar to m_τ.

There remains the question of physical equivalence among these representations. We shall show first that $V^{\tau,s}$ and $V^{-\tau,s}$ are anti-unitarily equivalent. Consider any anti-unitary isomorphism J of the Hilbert space $\mathscr{K}$ onto itself and define Θ by

$$(\Theta f)(\mathbf{p}) = Jf(-\mathbf{p}). \tag{253}$$

Θ is an anti-unitary isomorphism of $\mathscr{L}^2(\mathbf{P}^3,\mathscr{K},d\mathbf{p})$ onto itself. A quick calculation shows that

$$\Theta V^{\tau,s}\Theta^{-1} = V^{-\tau,s'},$$

where s' is the representation $h \to Js(h)J^{-1}$ of K^*. Since $\dim(s')=\dim(s)$, $s'\simeq s$, so that $V^{-\tau,s'}\simeq V^{-\tau,s}$. Hence $V^{\tau,s}$ and $V^{-\tau,s}$ are physically equivalent. Suppose conversely that V^{τ_1,s_1} and V^{τ_2,s_2} are physically equivalent. Let us first assume that the equivalence is effected by a *unitary* operator. Restriction to the subgroup of Euclidean motions ($\eta=0$, $\mathbf{v}=\mathbf{0}$) implies that the corresponding representations are unitarily equivalent and enables us to conclude, using lemma 9.13, that $s_1\simeq s_2$. We may (and do), therefore, assume that $s_1=s_2=s$ and that $V^{\tau_1,s}$ and $V^{\tau_2,s}$ both act in $\mathscr{H}=\mathscr{L}^2(\mathbf{P}^3,\mathscr{K},d\mathbf{p})$. It is then immediately obvious that the multipliers of the two representations $V^{\tau_1,s}$ and $V^{\tau_2,s}$ are similar. But we have observed that the multiplier of $V^{\tau,s}$ is similar to m_τ. So our observation implies that m_{τ_1} and m_{τ_2} are similar. The determination of the multiplier group of the Galilean group carried out in Chapter VII now enables us to conclude that τ_1 and τ_2 must be equal. Thus V^{τ_1,s_1} and V^{τ_2,s_2} cannot be unitarily physically equivalent unless $\tau_1=\tau_2$ and $s_1\simeq s_2$. If, on the other hand, they are anti-unitarily physically equivalent, then $V^{-\tau_1,s_1}$ and V^{τ_2,s_2} must be unitarily physically equivalent, and we could deduce that $\tau_2=-\tau_1$, $s_1\simeq s_2$. In other words, the projective representations $V^{\tau,s}$ ($\tau>0$) are, for distinct τ and inequivalent s, physically inequivalent, and any irreducible projective representation of the inhomogeneous Galilean group with nonexact multiplier is physically equivalent to one of these.

The action of $V^{\tau,s}$ can be described in another and perhaps more usual form by passing to Fourier-Plancherel transforms. Let (cf. (209))

$$\mathscr{F} : f \to \hat{f}$$

be the Fourier-Plancherel isomorphism of $\mathscr{L}^2(\mathbf{P}^3,\mathscr{K},d\mathbf{p})$ onto $\mathscr{L}^2(\mathbf{R}^3,\mathscr{K},d\mathbf{x})$ and let

$$\hat{V}^{\tau,s} = \mathscr{F} V^{\tau,s}\mathscr{F}^{-1}.$$

Then

$$\hat{V}_r^{\tau,s} = \exp\left(-\frac{i\eta}{2\tau}\nabla\right)\Gamma, \tag{254}$$

where

$$\nabla = \frac{\partial^2}{\partial x_1{}^2}+\frac{\partial^2}{\partial x_2{}^2}+\frac{\partial^2}{\partial x_3{}^2} \tag{255}$$

and

$$(\Gamma\hat{f})(\mathbf{x}) = \exp i\tau[\tfrac{1}{2}\langle\mathbf{u},\mathbf{v}\rangle+\langle\mathbf{x}-\mathbf{u},\, \mathbf{v}\rangle]s(h)\hat{f}(\delta(h)^{-1}(\mathbf{x}-\mathbf{u})). \tag{256}$$

∇ is self-adjoint with the usual domain. The expression (254) shows that the energy of the particle is the operator

$$E_\tau = -\frac{1}{2\tau}\nabla. \tag{257}$$

Notice that we have thus proved the formula (77) of Chapter VIII. Further, under the (covering group of the) Euclidean group, the states transform according to the law described by

$$(\hat{V}^{\tau,s}_{(h,0,\mathbf{0},\mathbf{u})}\hat{f})(\mathbf{x}) = s(h)\hat{f}(\delta(h)^{-1}(\mathbf{x}-\mathbf{u})). \tag{258}$$

A look at (258) is enough to conclude that the system corresponding to $V^{\tau,s}$ is localizable, and the projection valued measure which leads to the position operators is the usual one, $E \rightarrow P_E$, where

$$P_E : \hat{f} \rightarrow \chi_E \hat{f}. \tag{259}$$

In other words, $V^{\tau,s}$ describes the Schrödinger particle of *mass* τ and *spin* j where $\dim(s) = 2j+1$. We have thus obtained the following theorem:

Theorem 9.18. *The irreducible representations of the covering group of the inhomogeneous Galilean group G cannot represent nontrivial particles and the corresponding systems are not localizable. On the other hand, for each $\tau > 0$ and each half-integer $j = 0, \frac{1}{2}, 1, \cdots$ there is an irreducible projective representation with a nonexact multiplier of the inhomogeneous Galilean group, say $V^{\tau,j}$ and (with $s = D^j$) its action is given by* (251) *and* (252). *Every irreducible projective representation of G with a nonexact multiplier is physically equivalent to exactly one of the $V^{\tau,j}$. The particle corresponding to $V^{\tau,j}$ is the Schrödinger particle of mass τ and spin j, and its energy operator is $-(1/2\tau)\nabla$ where ∇ is the Laplacian* (255).

The analysis of the ordinary representations of the inhomogeneous Galilean group was first carried out by Inonu and Wigner [1], [2]. For the projective representations, the reader may consult the papers of Bargmann [1] and Wightman [1] (cf. also Pauli [1]).

NOTES ON CHAPTER IX

The mathematical and physical aspects of the theory of relativisitically invariant wave equation have spanned a vast literature; see, for instance, *Invariant Wave Equations*, Lecture Notes in Physics, No. 73, Springer-Verlag, Berlin, edited by G. Velo and A. S. Wightman, 1978. For an account of the early literature on this subject see E. M. Corson, *Introduction to Tensors, Spinors and Relativistic Wave Equations*, Chelsea reprint of the 1953 edition); see also Chapter IV of M. A. Naĭmark, *Linear Representations of the Lorentz Group* (in Russian), Moscow, 1958.

BIBLIOGRAPHY

Baer, R.

[1] *Linear Algebra and Projective Geometry*, Academic Press Inc., New York, 1952.

Bargmann, V.

[1] *On unitary ray representations of continuous groups*, Annals of Mathematics, **59** (1954), pp. 1–46.

[2] *Note on Wigner's theorem on symmetry operations*, Journal of Mathematical Physics, **5** (1964), pp. 862–868.

Bargmann, V., and Wigner, E. P.

[1] *Group theoretical discussion of relativistic wave equations*, Proceedings of the National Academy of Sciences, U.S.A., **34** (1948), pp. 211–223.

Birkhoff, G.

[1] *Lattice Theory*, American Mathematical Society Colloquium Publications, Vol. 25, 1960.

Birkhoff, G., and von Neumann, J.

[1] *The logic of quantum mechanics*, Annals of Mathematics, **37** (1936), pp. 823–843.

Blackwell, D.

[1] *On a class of probability spaces*, Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability, Vol. II (1956), pp. 1–6.

Bochner, S., and Chandrasekharan, K.

[1] *Fourier Transforms*, Princeton University Press, Princeton, 1949.

Bourbaki, N.

[1] *Eléments de Mathématique*, I, *Livre II*, *Algèbre*, Chs. 4 and 5, Hermann, Paris, 1959.

[2] *Eléments de Mathématique*, I, *Livre V*, *Espaces vectoriels topologiques*, Chs. 1 and 2, Hermann, Paris, 1953.

[3] *Eléments de Mathématique*, I, *Livre III*, *Topologie Générale*, Chs. 1 and 2, Hermann, Paris, 1961.

Brauer, R., and Weyl, H.

[1] *Spinors in n dimensions*, American Journal of Mathematics, **57** (1935), pp. 425–449.

Calabi, L.

[1] *Sur les extensions des groupes topologiques*, Annali de Matematica Pura ed Applicata, **32** (1951), pp. 295–370.

Cartan, E.

[1] *Leçons sur la théorie des spineurs*, I, II, Hermann, Paris, 1938.

Chevalley, C.

[1] *Theory of Lie Groups*, Princeton University Press, Princeton, 1946.

[2] *Théorie des Groupes de Lie*, Tome II, Hermann, Paris, 1951.

[3] *Théorie des Groupes de Lie*, Tome III, Hermann, Paris, 1955.

[4] *The Algebraic Theory of Spinors*, Columbia University Press, New York, 1954.

Chevalley, C., and Eilenberg, S.

[1] *Cohomology theory of Lie groups and Lie algebras*, Transactions of the American Mathematical Society, **63** (1948), pp. 85–124.

DIRAC, P. A. M.

[1] *The Principles of Quantum Mechanics*, Oxford University Press, London, 1958.

[2] *The quantum theory of the electron*, Proceedings of the Royal Society (London), **A 117** (1928), pp. 610–624.

DIXMIER, J.

[1] *Les algebres d'opérateurs dans l'espace Hilbertien*, Gauthier-Villars, Paris, 1957.

DUNFORD, N., AND SCHWARTZ, J.

[1] *Linear Operators*, Part 1, Interscience Publishers, Inc., New York, 1958.

[2] *Linear Operators*, Part 2, Interscience Publishers, Inc., New York, 1963.

EILENBERG, S.

[1] *Topological methods in abstract algebra. Cohomology theory of groups*, Bulletin of the American Mathematical Society, **55** (1949), pp. 3–37.

FEDERER H., AND MORSE, A. P.

[1] *Some properties of measurable functions*, Bulletin of the American Mathematical Society, **49** (1943), pp. 270–277.

FEYNMAN, R. P.

[1] *Concept of probability in quantum mechanics*, Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability, Berkeley, (1951), pp. 533–541.

[2] *Space-time approach to non-relativistic quantum mechanics*, Reviews of Modern Physics, **20** (1948), pp. 367–387.

FINKELSTEIN, D., JAUCH, J. M., SCHMINOVICH, S., AND SPEISER, D.

[1] *Foundations of quaternion quantum mechanics*, Journal of Mathematical Physics, **3** (1962), pp. 207–220.

GEL'FAND, I. M.

[1] *On one-parametrical groups of operators in a normed space*, Doklady Akademii Nauk, SSSR, **25** (1939), pp. 713–718.

GEL'FAND, I. M., AND NAĬMARK, M. A.

[1] *Unitary representations of the group of linear transformations of the straight line*, Doklady Akademii Nauk, SSSR, **55** (1947), pp. 567–570.

GLEASON, A. M.

[1] *Measures on the closed subspaces of a Hilbert space*, Journal of Mathematics and Mechanics, **6** (1957), pp. 885–894.

GUDDER, S. P.

[1] *Spectral methods for a generalized probability theory*, Transactions of the American Mathematical Society, **119** (1965), pp. 428–442.

HALMOS, P. R.

[1] *Measure Theory*, D. van Nostrand Company, Princeton, 1950.

[2] *Introduction to the Theory of Hilbert Space and Spectral Multiplicity*, Chelsea Publishing Co., New York, 1957.

HEISENBERG, W.

[1] *The Physical Principles of Quantum Theory*, Dover Publications, Inc., New York, 1930.

HELGASON, S.

[1] *Differential Geometry and Symmetric Spaces*, Academic Press Inc., New York, 1962.

HELSON, H.

[1] *Lectures on Invariant Subspaces*, Academic Press Inc., New York, 1964.

HELSON, H., AND LOWDENSLAGER, D.

[1] *Invariant subspaces*, Proceedings of the International Symposium on Linear Spaces, pp. 251–262 (Jerusalem, 1960), Pergamon Press, Inc., Oxford, 1961.

Heyting, A.
[1] *Axiomatic Projective Geometry*, Interscience Publishers, Inc., New York; P. Noordhoff N. V., Groningen, North Holland Publishing Company, Amsterdam, 1963.

Hochschild, G.
[1] *Group extensions of Lie Groups*, I, Annals of Mathematics, **54** (1951), pp. 96–109.
[2] *Group extensions of Lie Groups*, II, Annals of Mathematics, **54** (1951), pp. 537–551.

Inonu, E., and Wigner, E. P.
[1] *Representations of the Galilei group*, Il Nuovo Cimento, **9** (1952), pp. 705–718.
[2] *On the contraction of groups and their representations*, Proceedings of the National Academy of Sciences, U.S.A., **39** (1953), pp. 510–524.

Jacobson, N.
[1] *Lectures in Abstract Algebra*, Vol. II, D. Van Nostrand Company, Inc., Princeton, 1953.
[2] *Lie Algebras*, Interscience Publishers, Inc., New York, 1962.

Jordan, P., von Neumann, J., and Wigner, E. P.
[1] *On an algebraic generalization of the quantum mechanical formalism*, Annals of Mathematics, **35** (1934), pp. 29–64.

Kakutani, S., and Mackey, G. W.
[1] *Ring and lattice characterizations of complex Hilbert space*, Bulletin of the American Mathematical Society, **52** (1946), pp. 727–733.

Kaplansky, I.
[1] *Any orthocomplemented complete modular lattice is a continuous geometry*, Annals of Mathematics, **61** (1955), pp. 524–541.
[2] *Infinite Abelian Groups*, University of Michigan Press, Ann Arbor, 1954.

Kelley, J. L.
[1] *General Topology*, D. Van Nostrand Company, Princeton, 1955.

Kleppner, A.
[1] *Multipliers on abelian groups*, Mathematische Annalen, **158** (1965), pp. 11–34.

Kolmogorov, A. N.
[1] *Zur Begründung der Projectiven Geometrie*, Annals of Mathematics, **33** (1932), pp. 175–176.

Koopman, B. O.
[1] *Hamilton systems and transformations in Hilbert space*, Proceedings of the National Academy of Sciences, U.S.A., **17** (1931), pp. 315–318.

Kuratowski, C.
[1] *Topologie*, Vol. I, Warsaw, 1948.
[2] *Topologie*, Vol. II, Warsaw, 1950.

Landau, L. D., and Lifchitz, E. M.
[1] *Quantum Mechanics, non-relativistic Theory*, Gos. Izd. Phys. Mat. Moscow, 1963.

Loomis, L. H.
[1] *On the representation of σ-complete Boolean algebras*, Bulletin of the American Mathematical Society, **53** (1947), pp. 757–760.
[2] *An Introduction to Abstract Harmonic Analysis*, D. Van Nostrand Company, Inc., Princeton, 1953.
[3] *The lattice theoretic background of the dimension theory of operator algebras*, Memoirs of the American Mathematical Society, **18** (1955), pp. 1–36.

Mackey, G. W.
[1] *The Mathematical Foundations of Quantum Mechanics*, W. A. Benjamin, Inc., New York, 1963.
[2] *Imprimitivity for representations of locally compact groups*, Proceedings of the National Academy of Sciences, U.S.A., **35** (1949), pp. 537–545.

[3] *Induced representations of locally compact groups*, I, Annals of Mathematics, **55** (1952), pp. 101–139.

[4] *Induced representations of locally compact groups*, II, Annals of Mathematics, **58** (1953), pp. 193–221.

[5] *Borel structure in groups and their duals*, Transactions of the American Mathematical Society, **85** (1957), pp. 134–165.

[6] *Unitary representations of group extensions*, I, Acta Mathematica, **99** (1958), pp. 265–311.

[7] *Infinite dimensional group representations*, Bulletin of the American Mathematical Society, **69** (1963), pp. 628–686.

[8] *Les ensembles Boreliens et les extensions des groupes*, Journal de Mathématiques Pures et Appliquées, **36** (1957), pp. 171–178.

[9] *A theorem of Stone and von Neumann*, Duke Mathematical Journal, **16** (1949), pp. 313–326.

MALČEV, A.

[1] *On the simple connectedness of invariant subgroups of Lie groups*, Doklady Akademii Nauk, SSSR, **34** (1942), pp. 10–13.

MAUTNER, F. I.

[1] *Unitary representations of locally compact groups*. I, Annals of Mathematics, **51** (1950), pp. 1–25.

[2] *Unitary representations of locally compact groups*. II, Annals of Mathematics, **52** (1950), pp. 528–556.

MONTGOMERY, D., AND ZIPPIN, L.

[1] *Topological transformation groups*, Interscience Publishers, Inc., New York, 1955.

MOORE, C. C.

[1] *Extensions and low dimensional cohomology theory of locally compact groups*, I, Transactions of the American Mathematical Society, **113** (1964), pp. 40–63.

MOYAL, J. E.

[1] *Quantum Mechanics as a statistical theory*, Proceedings of the Cambridge Philosophical Society, **45** (1949), pp. 99–124.

MURRAY, F. J., AND VON NEUMANN, J.

[1] *On rings of operators*, Annals of Mathematics, **37** (1936), pp. 116–229.

[2] *On rings of operators*, II, Transactions of the American Mathematical Society, **41** (1937), pp. 208–248.

[3] *On rings of operators*, IV, Annals of Mathematics, **44** (1943), pp. 716–808.

NAKANO, H.

[1] *Unitärinvariante hypermaximale normale operatoren*, Annals of Mathematics, **42** (1941), pp. 657–664.

NEWTON, T. D., AND WIGNER, E. P.

[1] *Localized states for elementary systems*, Reviews of Modern Physics, **21** (1949), pp. 400–406.

PAULI, W.

[1] *Die allgemeinen Prinzipen der Wellenmechanik*, Handbuch der Physik, 2 Aufl. Band **24**, 1 Teil (Reprint, Edwards Brothers, Inc., Ann Arbor, Michigan), 1950.

PIRON, C.

[1] *Axiomatique Quantique*, Thèse, Université de Lausanne, Faculté des Sciences, Bale, Imprimerie Birkhauser S. A. (1964), pp. 439–468.

PONTRJAGIN, L. S.

[1] *Topological Groups*, Princeton University Press, Princeton, 1939.

RIESZ, F., AND NAGY, B. SZ.

[1] *Leçons d'analyse Fonctionnelle*, Akadémiai Kiadó, Budapest, 1952.

SCHWARTZ, L.

[1] *Théorie des distributions*, Tome I, Hermann, Paris, 1957.

[2] *Théorie des distributions*, Tome II, Hermann, Paris, 1959.

[3] *Application of distributions to the study of elementary particles in relativistic quantum mechanics*, Technical Report No. 7, Department of Mathematics, University of California, Berkeley, 1961, pp. 1–207.

SCHWEBER, S.

[1] *An introduction to relativistic quantum field theory*, Row, Peterson and Company, Evanston and New York, 1961.

SEGAL, I. E.

[1] *Postulates for general Quantum Mechanics*, Annals of Mathematics, **48** (1947), pp. 930–948.

[2] *Quantization of nonlinear systems*, Journal of Mathematical Physics, **1** (1960), pp. 468–488.

[3] *Mathematical problems of relativistic physics*, Lectures in Applied Mathematics, American Mathematical Society, Providence, R.I., 1963.

SEIDENBERG, A.

[1] *Lectures in Projective Geometry*, D. Van Nostrand Company, Inc., Princeton, 1962.

SIKORSKI, R.

[1] *Boolean Algebras*, Springer, Berlin, 1964.

STERNBERG, S.

[1] *Lectures on Differential Geometry*, Prentice-Hall, Inc., Englewood Cliffs, N.J., 1964.

STONE, M. H.

[1] *The theory of representation for Boolean algebras*, Transactions of the American Mathematical Society, **40** (1936), pp. 37–111.

[2] *Linear transformations in Hilbert space and their applications to analysis*, American Mathematical Society Colloquium Publications, Vol. 15, 1951.

[3] *Linear transformations in Hilbert space.* III. *Operational methods and group theory*, Proceedings of the National Academy of Sciences, U.S.A., **16** (1930) pp. 172–175.

VAN HOVE, L.

[1] *Sur certaines représentations unitaires d'un groupe infini de transformations*, Academie Royale de Belgique Classe des Sciences, Memoires Collection in 8°, **29**, No. 6 (1951), pp. 1–102.

VARADARAJAN, V. S.

[1] *Probability in physics and a theorem on simultaneous observability*, Communications in Pure and Applied Mathematics, **15** (1962), pp. 189–217; correction, *loc. cit.*, **18** (1965).

[2] *Groups of automorphisms of Borel spaces*, Transactions of the American Mathematical Society, **109** (1963), pp. 191–220.

[3] *Measures on topological spaces*, Matematiceskiĭ Sbornik, **55** (97) (1961), pp. 35–100.

[4] *Special topics in probability*, mimeographed notes of lectures delivered at the Courant Institute of Mathematical Sciences, New York University, 1962.

[5] *Global theory of representations of locally compact groups*, (Notes by Alan Solomon)—mimeographed notes of lectures delivered at the Courant Institute of Mathematical Sciences, New York University, 1961–62.

VON NEUMANN, J.

[1] *Mathematical Foundations of Quantum Mechanics*, Princeton University Press, Princeton, 1955 (translated by Robert T. Beyer from Mathematische Grundlagen der Quantenmechanik, Springer, Berlin, 1932).

[2] *On rings of operators. Reduction theory*, Annals of Mathematics, **50** (1949), pp. 401–485.

[3] *Die Eindeutigkeit der Schrödingerschen Operatoren*, Mathematische Annalen, **104** (1931), pp. 570–578.

[4] *On an algebraic generalisation of the quantum mechanical formalism* (part I), Rec. Math. Matematiceskiĭ Sbornik, N.S., **1** (1936), pp. 415–484.

[5] *Thermodynamik quantenmechanischer Gesamtheiten*, Göttinger Nachrichten (1927), pp. 273–291.

[6] *Beweis des Ergodensatzes und des H-Theorems in der neuen Mechanik*, Zeitschrift für Physik, **57** (1929), pp. 30–70.

WEIL, A.

[1] *L'integration dans les groupes topologiques et ses applications*, Hermann, Paris, 1938.

WEISS, E., AND ZIERLER, N.

[1] *Locally compact division rings*, Pacific Journal of Mathematics, **8** (1958), pp. 369–371.

WEYL, H.

[1] *The Theory of Groups and Quantum Mechanics*, Dover Publications, Inc., New York, 1931 (translated by H. P. Robertson from Gruppentheorie und Quantenmechanik).

[2] *The Classical Groups*, Princeton University Press, Princeton, 1946.

[3] *Theorie der Darstellung kontinuerlicher halb-einfacher Gruppen durch lineare Transformationen*, Teil I, Mathematische Zeitschrift, **23** (1925), pp. 271–309.

[4] ———, Teil II, Mathematische Zeitschrift, **24** (1926), pp. 328–376.

[5] ———, Teil III, Mathematische Zeitschrift, **24** (1926), pp. 377–395.

[6] *Space, Time, Matter*, Dover Publications, Inc., 1950 (translated by Henry L. Brose from Raum, Zeit, Materie, 1921).

WICK, G. C., WIGHTMAN, A. S., AND WIGNER, E. P.

[1] *The intrinsic parity of elementary particles*, The Physical Review, **88** (1952), pp. 101–105.

WIGHTMAN, A. S.

[1] *On the localizability of quantum mechanical systems*, Reviews of Modern Physics, **34** (1962), pp. 845–872.

[2] *Relativistic Invariance and Quantum Mechanics*, Supplemento al Vol. XIV, serie X, del Nuovo Cimento (1959), pp. 81–94 (Notes by A. Barut).

WIGNER, E. P.

[1] *Unitary representations of the inhomogeneous Lorentz group*, Annals of Mathematics, **40** (1939), pp. 149–204.

[2] *On the quantum correction for thermodynamic equilibrium*, The Physical Review, **40** (1932), pp. 749–759.

[3] *Group theory and its applications to the quantum mechanics of atomic spectra*, Academic Press Inc., New York, 1959.

[4] *Über die Operation der Zeitumkehr in der Quantenmechanik*, Göttinger Nachrichten Math-Physik Kl (1931), pp. 546–559.

[5] *Relativistische Wellengleichungen*, Zeitschrift für Physik, **124** (1947), pp. 665–684.

[6] *Relativistic invariance and quantum phenomena*, Reviews of Modern Physics, **29** (1957), pp. 255–268.

ZIERLER, N.

[1] *Axioms for non-relativistic quantum mechanics*, Pacific Journal of Mathematics, **11** (1961), pp. 1151–1169.

[2] *On the lattice of closed subspaces of Hilbert space* (forthcoming).

INDEX

Absolute continuity (of measures), 151
Affine coordinate system, 301
Affine space, 160, 300
  automorphism of, 300
  translation in, 300
  vector space associated with, 300
Angular momentum
  around an axis, 310
  orbital, 311, 364
  spin, 311, 364
Annihilation operator, 317
Annihilator (of a linear manifold), 23
Anti-automorphism (of a division ring), 23

Banach space, over a division ring, 73
Basis (in a generalized geometry), 33
Bell's inequality (for local hidden variable models), 131
Θ-bilinear form, 23
  definite, 26
  right/left nonsingularity of, 23
  symmetric, 24
Boolean algebra, 8
Boolean σ-algebra, 8
Boolean subalgebra of a logic, 55
Boolean sub σ-algebra of a logic, 55
Borel automorphism, 148
Borel G-space, 159
  G-subspace of, 159
  standard, 159
Borel group, 157
  separable, 157
  standard, 157
Borel isomorphism, 148
Borel mapping, 148
Borel section, 164
  regular, 164
Borel space, 148
  separable, 149
  standard, 150
  subspace of, 149
Borel structure, 148
  natural (of a topological space), 149
Bose–Einstein statistics, 139, 317
Bosons, 139, 318
Bounded set (in a topological group), 196

C*-algebra, 134
Center (of a complemented modular lattice), 19
Center (of a logic), 64
  atoms of, 64
  continuity of, 64
Chain (in a lattice), 18
  length of, 18
Clifford algebra, 350
  concrete, 348
Closed form, 273
Coboundary, 175
  strict, 175
Cocycle
  relative to a measure class, 174
  strict, 174

Coherent subspace of states, 114
Cohomologous cocycles, 174
Cohomology class, 175
  strict, 175
Cohomology group, 175
Collinearity (of points in a lattice), 20
Commutation rules, 5, 294
  of Heisenberg, 294
Commuting ring (of a representation), 203
Complement (of an element in a lattice), 8
Complementarity, 302
Completion (of a σ-algebra with respect to a measure), 151
Configuration observables, 5, 295
Configuration space, 4, 293
Continuous action (of a lcsc group), 159
Continuous geometries with a transition probability, 71
Continuous section, 164
Coordinates, in a generalized geometry, 28
  affine, 34
  homogeneous, 37
Coordinatization
  of a geometry, 20
  of a generalized geometry, 38
  of a complemented modular lattice, 40
Creation operator, 317
Cross-section lemma (of von Neumann), 154

Degrees of freedom (of a mechanical system), 1
Desarguesian generalized geometry, 29
Desarguesian planes, 28
Dimension (of an element), 19, 29
Dimension function (of a lattice), 19
Dirac equation (of the electron), 369
Dirac γ-matrices, 355
Dispersion free ensembles, 123
Dispersion free states, 124
Dual of a division ring, 22
Dual ideal, 9
Duality (between geometries), 23
Dynamical group
  of a classical system, 6
  of a general quantum system, 68
Dynamical states, 292, 327

Einstein–Podolsky–Rosen paradox, 129, 137
Energy, 290
Entropy
  macroscopic, 142
  of a quantum mechanical ensemble, 142
  of a quantum mechanical state, 141
Ergodic measure class, 182
Euclidean affine space, 303
  motion of, 304
Euclidean coordinate system, 304
Exact form, 273
Extension (of one lcsc group by another), 251
  central, 251
  equivalence of, 251
Extension, central, of one Lie algebra by another, 267
  associated with a group extension, 268
  covering of one by another, 269
  equivalence of, 268
Exterior algebra over a Hilbert space, 139, 319

Fermi–Dirac statistics, 139, 317
Fermions, 139, 319
Frame (of a generalized geometry) at a point, 33
Frame function, 85
  regular, 85
  weight of, 85
Free action of a group, 183

Free particle, 69
  classical, 322
  quantum, 330, 399

G-Hilbert space bundle, 234
  section of, 234
  square integrable section of, 234
G-homomorphism, 159
G-isomorphism, 159
G-space, 158
  homogeneous, 160
  locally compact, 159
  transitive, 160
Galilean group, 283, 325
  projective representation of, 394
Galilean relativity, 325, 391
Galilean transformation, 283
Gauge condition, 371
Generating sequence (for a Borel structure), 149
Geometry, 19
  generalized, 29
  projective, 18
Gleason's theorem, 97
  generalizations of, to von Neumann algebras, 146

Haar measure, 157
Hamiltonian, 1, 3, 290
Heisenberg picture, 293
Heisenberg states, 293, 327
Hidden variables in Quantum Mechanics, 122, 131
Hilbert space over a division ring, 26
Hilbertian pairs, 115
Homomorphism
  between Boolean algebras, 9
  defined by a cocycle at a point, 176
  equivalence of, 176
σ-homomorphism between Boolean σ-algebras, 10

Independence (of points), 19, 32
Inertial coordinate system, 323
Inertial observer, 323
Infinitesimal operators (for the free particle representations), 364
Infinity
  element at, 33
  points at, lines at, etc., 33
Inhomogeneous group, associated with a group and a representation of it, 278
Integral of motion, 292
Inversions, group of, 337
Involutive anti-automorphism, 25
Isomorphism (between geometries), 21
  linearity of, 21

Jordan algebras, 133
  formally real, 133

Kernel of a σ-homomorphism, 11
Koopman system of imprimitivity, 208

lcsc group, 156
Lattice, 8
  complemented, 8
  distributive, 8
  irreducible, 19
  lines, planes, etc., of, 20
  modular, 18
  points of, 19
  rank of, 19
Linear momentum in a given direction, 301
Localizable systems, 383
Locally bounded map (from one topological space to another), 181
Locally bounded topological group, 196
Logic, 42
  associated with a division ring, 119

Logic (*cont'd.*)
  associated with a von Neumann algebra, 112
  of a classical system, 6
  injection of, 43
  irreducible, 64
  isomorphism of, 43
  projective, 114
  standard, 44, 74, 80
  sublogic of, 43
Lorentz condition, 371
Lorentz group, 326
Lorentz–Hermite form, 363

Maxwell equations, 371
Measure, 150
  Borel, 150
  finite, 150
  σ-finite, 150
  invariant, 151, 159
  quasi-invariant, 159
  regular, 150
  standard, 150
Measure class, 151
  bi-invariant, 191
  corresponding to the multiplicity *n*, 219
  ergodic, 182
  invariant, 151
  living on an orbit, 223
  of projection valued measure, 218, 219
  transitive, 182
Measure on a logic, 50
Minkowskian form, 326
Modular function (of a lcsc group), 157
Momentum
  angular, 4
  corresponding to symmetries, 4
  linear, 4
Momentum observables, 5, 295
Momentum space, 339
  representations, 329
Momentum vector, 1
Motions of a quantum mechanical system, 292
Multiplier (*K*-multiplier), 247
  exact, 248
  locally continuous, 255
  similar, 248
Multiplier group (*K*-multiplier group), 248
Multiplier of a projective representation, 248
Mutually singular measures, 218

Observables
  associated with a Boolean σ-algebra, 12
  associated with a logic, 46
  associated with a standard logic, 82
  bounded, 47
  of a classical system, 2, 3
  central, 64
  discrete, 47
  expected value of, 52
  probability distribution of, 52, 63
  spectrum of, 47
  value of, 47
  variance of, 52
Operators of trace class, 83
Orbit of a point, 159
Orthocomplementation
  in the geometry of a vector space, 26
  in a lattice, 42
Orthogonal complement (of an element), 43

Partial Boolean algebra, 127
Partially ordered set, 8
  null element of, 8
  unit element of, 8
Phase space, 2
Physical states of von Neumann algebras, 125, 146
Polarity, 24
  isotropic, 24
Position operators
  for nonrelativistic particles, 301, 302

- for relativistic particles, 377, 387, 390
Product of Borel spaces, 149
Projection valued measures, 80
  - corresponding to a representation of an abelian group, 238
  - cyclic vector for, 218
  - homogeneous, 218
  - measure class of, 218, 219
  - multiplicity of, 218

Projective group (of a Hilbert space), 244
Projective representation, 248
  - multiplier of, 248

Projectivity
  - canonical fixed points of, 39
  - general, 39
  - special, 39

Pure state, 53

Quantization, 315
Quantum probability, 145
  - conditional, 145

Quasi invariant measure, 159
Quasi-probability distribution, 104
Quaternions, 21
  - conjugacy of, 22
  - conjugation of, 22
  - norm of, 21
  - unit, 22

Question observables, 50

Radon–Nikodym derivative, 151
Range of an observable, 60
Regular ring, 40
Representation
  - absolutely irreducible, 278
  - admissible, 278
  - equivalence of, 202
  - of a group in the group of automorphisms of the convex set of states, 69
  - irreducible unitary, 202
  - physical equivalence of, 305
  - unitary, 201

Schrödinger equation (abstract), 291
Schrödinger particle, 399
Schrödinger picture, 293
Schrödinger representation (of commutation rules), 299
Second Quantization, 316, 321
Semidirect product, 236
Semilinear transformation, 20
Separable σ-algebras, 15, 53
Separating sequence for a Borel structure, 149
Simultaneous observability, 55
  - verifiability, 55

Smooth orbit structure, 228
Space inversion, 337
Space-time inversion, 337
Space-time relativity group, 323
Spectrum of a relativistic system, 329
Spin (of a free relativistic particle), 348
Spin matrices (of Pauli), 307
Spin observables, 311
Spin representation, 353
Spinor field, 356
Spinors, 353
Stability subgroup at a point, 159
State
  - of a classical system, 1
  - of a logic, 52
  - of a standard logic, 85

State function (of a logic), 49
States, coherent subspace of, 114
Statistical interpretation of Born, 71
Strict coboundary, 175
Strict cocycle, 174
Strict cohomology class, 175
Strictly cohomologous strict cocycles, 175
Stone space, 9
Stone topology, 10
Strong topology (for the unitary group of a Hilbert space), 193
Subrepresentation, 204
Superposition of states
  - classical, 53
  - of a logic, 53

Superposition principle, 54, 98

Superselection rules, 114
Symmetric algebra over a Hilbert space, 139, 318
Symmetry
  of a Hilbert space, 104
  physical, 108
  Θ-unitary, 104
Systems of imprimitivity for a group, 203
  associated with a cocycle and a measure class, 217
  direct sum of, 203
  equivalence of, 203
  induced by a representation of a subgroup, 223
  irreducible, 203
  living on an orbit, 223
  physical equivalence of, 305
  transitive, 223
Time inversion, 325, 337
Total energy, 329
Transition probability (between states), 107

Uncertainty principle, 66, 71
Unimodular group, 157
Unitary transformation, 28

Vacuum, 317, 347
Velocity observables, 391
von Neumann algebras, 112
von Neumann operators, 85

Wave equations, 365, 399
  of particles, 369, 370, 371
Wave function, 302
Weil topology, 195

This is a book on the foundations of Quantum Theory. It examines the subject from a point of view that goes back to Heisenberg and Dirac, and whose definitive mathematical formulation is due to von Neumann. This view, which is essentially geometric, leads most naturally to the questions of symmetry that are at the basis of all attempts to understand the world of atoms and subatomic particles. The mathematical treatment of symmetry in Quantum Theory is based on the theory of group representations, and this book includes a self-contained treatment of the parts of this theory which are among the most useful in Quantum Physics. This edition, the first Springer edition of the previous two-volume work, has been revised to bring it up to date.

ISBN 0-387-**96124**-0
ISBN 3-540-**96124**-0